

/translation: "MASTIKALSVSEFOSLEFCATVSPILATENTASSSEYECOT
SKMSRPVQODWLSQPPARVITIKMECNQVNGSRNSDDSCVAKGMSVSDNVG
NYSYMEKHIIPNNMTINERVIVPADPLMSDHWQMEWAKVEGLDVLILF
ONIDKEIKMTKDDFORLTSTYNADILSLHLREPLHSIDVDKALQNSPRL
MHANRGATFIENPNTSYRPAOTIRTRPDLPEARSPAMTSISHSTOKAOPS
STVKTDORPOLRPOILOIGPTSSRIANPGSGOIOLOFLELSDSNSTCITMEGT
NGEKTMDPDEVARMRERKSPNNNTDKSLRALRYITDKINIKVHKKRRAIFEDRI
GIALOPHPRESSMKIPSDLPYMSSTHAPQKNEVAPHPALPVTSSSFAAPNP
YMSPTGITYNTRLPAAHMPSHLGYT"

BASE COUNT 458 a 392 c 327 g 339 t
ORIGIN

Query Match 65.68; Score 949; DB 21; Length 1516;
Best Local Similarity 94.78; Pred. No. 0.00e+00;
Matches 1435; Conservative 0; Mismatches 0; Indels 81; Gaps 1;

Db 1 GAATCCCGGAAAGATATTTATTTAGCAATTTAGCGATCATATATCTTGATCA 60
1 gaattcccggaagataattattatlaagcaattlaagcattcaatacttgatcaca 60
Oy 61 TTATGCAAGCACTTTAAGAGCATTTATCAGTGTGAGTGAAGACCAAGTCTTTG 120
61 ttatgcaagcactttaagagcatattatcagtggtgagagaacagctcttggtg 120
Oy 61 AGTGCGCTACGAGATCGCCCACTTGCAAAAGACAAATGACAGCTCTCTTCCAGTG 180
121 agtgcgctacgagatcgcccaacttgcaaaagacagaatgacagctctcttcagtg 180
Oy 121 AATATGGGCAAAATCAAGATGAGCCGCGGTCCCGAGAGACTGGTATACAGAC 240
121 aatatgggcaaaatcaagatgagccgcggtcccgagagactgggtatcacagac 240
Oy 121 AATATGGGCAAAATCAAGATGAGCCGCGGTCCCGAGAGACTGGTATACAGAC 240
121 aatatgggcaaaatcaagatgagccgcggtcccgagagactgggtatcacagac 240
Oy 241 CCCCAGCAGAGATTACATTAAGATGAGTGTACCAACCAACAGATTAGGGTCAAGGA 300
241 cccagcagagattacatthaagatgagtgtaaccacaacaggttaatggtcaagga 300
Oy 241 CCCCAGCAGAGATTACATTAAGATGAGTGTACCAACCAACAGATTAGGGTCAAGGA 300
241 cccagcagagattacatthaagatgagtgtaaccacaacaggttaatggtcaagga 300
Oy 301 ATTCACTGATGACTGCAACCGTGGCAAAAGAGGAAATGGTTAGCAGTTGACACATG 360
301 attcactgatgactgcaaccgtggcaaaagaggaaatggttagcagttcagaacatg 360
Oy 301 attcactgatgactgcaaccgtggcaaaagaggaaatggttagcagttcagaacatg 360
301 attcactgatgactgcaaccgtggcaaaagaggaaatggttagcagttcagaacatg 360
Oy 361 TTGGATGAACATGAGAGCTACATGGAAGAGATGATGCTTCCAGAGCTGACATCTTGTGT 540
361 ttggatgaacatgagagctacatggaaagatgattgcttccagagctgacatcttgtgt 540
Oy 361 ttggatgaacatgagagctacatggaaagatgattgcttccagagctgacatcttgtgt 540
361 ttggatgaacatgagagctacatggaaagatgattgcttccagagctgacatcttgtgt 540
Oy 421 CCAATGAACGAAGATTATTTGTCCAGCAGATCTACGTTATGAGACAGACATGTAC 480
421 ccaatgaacgaagattatttgtccagcagatctacgttatgagacagacatgtac 480
Oy 421 ccaatgaacgaagattatttgtccagcagatctacgttatgagacagacatgtac 480
421 ccaatgaacgaagattatttgtccagcagatctacgttatgagacagacatgtac 480
Oy 481 GGCAGTGGCTGGAGTGGGAGTGAAGAGTATGCTTCCAGAGCTGACATCTTGTGT 540
481 ggcagtggctggagtgggagtgaaagtattgcttccagagctgacatcttgtgt 540
Oy 481 ggcagtggctggagtgggagtgaaagtattgcttccagagctgacatcttgtgt 540
481 ggcagtggctggagtgggagtgaaagtattgcttccagagctgacatcttgtgt 540
Oy 541 TCCACAAGATTGATGGGAAAGTGTGTAATAACCAAGATGACTTCCAGACATCA 600
541 tccacaagattgatgggaaagtgtgtaataaccagaatgacttccagacatca 600
Oy 541 tccacaagattgatgggaaagtgtgtaataaccagaatgacttccagacatca 600
541 tccacaagattgatgggaaagtgtgtaataaccagaatgacttccagacatca 600
Oy 601 CGCCAGGCTATTAACGACATATCTCTGTCACTACACTACACTACAGAGATCTCTC 660
601 cgccaggcttatTAACGACATATCTCTGTCACTACACTACACTACAGAGATCTCTC 660
Oy 601 cgccaggcttatTAACGACATATCTCTGTCACTACACTACACTACAGAGATCTCTC 660
601 cgccaggcttatTAACGACATATCTCTGTCACTACACTACACTACAGAGATCTCTC 660
Oy 661 TTCCACATTTGACTTACATGATGTTGATAGGCTTACAAAACCTCTCCAGGTTAATGC 720
661 ttccacatTTGACTTACATGATGTTGATAGGCTTACAAAACCTCTCCAGGTTAATGC 720
Oy 661 ttccacatTTGACTTACATGATGTTGATAGGCTTACAAAACCTCTCCAGGTTAATGC 720
661 ttccacatTTGACTTACATGATGTTGATAGGCTTACAAAACCTCTCCAGGTTAATGC 720
Oy 721 ATGCTAGAAACACAGAGAGCCACTTTTATTTTCCAAATACATCAGTTTACCAGAG 780
721 atgctagaaacacagagagccactTTTATTTTCCAAATACATCAGTTTACCAGAG 780
Oy 721 atgctagaaacacagagagccactTTTATTTTCCAAATACATCAGTTTACCAGAG 780
721 atgctagaaacacagagagccactTTTATTTTCCAAATACATCAGTTTACCAGAG 780
Oy 781 CAACGCAAAAGATTAACAAAGCCAGATTTTACCTTATGAGCAGAGAGATCAGCGT 840
781 caacgcaaaagattAAACAAAGCCAGATTTTACCTTATGAGCAGAGAGATCAGCGT 840
Oy 781 caacgcaaaagattAAACAAAGCCAGATTTTACCTTATGAGCAGAGAGATCAGCGT 840
781 caacgcaaaagattAAACAAAGCCAGATTTTACCTTATGAGCAGAGAGATCAGCGT 840
Oy 700 caacgcaaaagattAAACAAAGCCAGATTTTACCTTATGAGCAGAGAGATCAGCGT 759
700 caacgcaaaagattAAACAAAGCCAGATTTTACCTTATGAGCAGAGAGATCAGCGT 759
Oy 700 caacgcaaaagattAAACAAAGCCAGATTTTACCTTATGAGCAGAGAGATCAGCGT 759
700 caacgcaaaagattAAACAAAGCCAGATTTTACCTTATGAGCAGAGAGATCAGCGT 759
Oy 841 GGACGAGTCACAGCCATCCACTCAGTCAAAAGGTAACCAACATCATCTTCAACAGTGC 900
841 ggacgagtcacagccatccactcagtcAAAAGGTAACCAACATCATCTTCAACAGTGC 900

Oy 760 ggaagagtcacagccatcccaactcaagtcataaagctaccacacatcattcaacagtc 819
Oy 901 CCAAAACGAAGACAGCGGCTCTCAGTAGATCCCTTATCAATCTTGGACGACACCA 960
901 ccaaaacgaagacagcgctctcagtagatcccttatacagatctcttgaccagcaga 879
Oy 820 ccaaaacgaagacagcgctctcagtagatcccttatacagatctcttgaccagcaga 879
820 ccaaaacgaagacagcgctctcagtagatcccttatacagatctcttgaccagcaga 879
Oy 961 GCCGCTTTCGAATTCAGGAGTGGGAGATACAGCTATGCGACTTCTTACTGAGCTTC 1020
961 gccgcttTCGAATTCAGGAGTGGGAGATACAGCTATGCGACTTCTTACTGAGCTTC 1020
Oy 880 gccgcttTCGAATTCAGGAGTGGGAGATACAGCTATGCGACTTCTTACTGAGCTTC 939
880 gccgcttTCGAATTCAGGAGTGGGAGATACAGCTATGCGACTTCTTACTGAGCTTC 939
Oy 1021 TGTGGAGAGCTCCCACTCCCACTCAGTACCTGAGAGGACCAAAATGGGAGTTCAAGA 1080
1021 tgtggagagctcccaactcccaactcagtagctgagagggacaaaatgggagttcaaga 1080
Oy 940 tgcggaagctcccaactcccaactcagtagctgagagggacaaaatgggagttcaaga 999
940 tgcggaagctcccaactcccaactcagtagctgagagggacaaaatgggagttcaaga 999
Oy 1081 TGACAGACCTCGATGAATGCGTGGCGCTTGGGAGACAGGAAAAGCAAACTTAACATGA 1140
1081 tgacagacctcgatgaatgcggtggcgcttgggagacagggaaaagcaaaacttaacatga 1140
Oy 1000 tgaagacctcgatgaatgcggtggcgcttgggagagagagaaagcaaaacttaacatga 1039
1000 tgaagacctcgatgaatgcggtggcgcttgggagagagagaaagcaaaacttaacatga 1039
Oy 1141 ACTATGACAAATCAGCGGCTGCACTTCCGCTACTACTATGACAAAATTTATGACTAAG 1200
1141 actatgacaaatcagcggtgcaacttccgctactactatgacaaaatTTATGACTAAG 1200
Oy 1060 actatgacaaatcagcggtgcaacttccgctactactatgacaaaatTTATGACTAAG 1119
1060 actatgacaaatcagcggtgcaacttccgctactactatgacaaaatTTATGACTAAG 1119
Oy 1201 TTGATGTTAAACGCTATGCTTCAAAATTTGATTTTCACAGGAATGCTAGGCTCCACG 1260
1201 ttgatgttAAACGCTATGCTTCAAAATTTGATTTTCACAGGAATGCTAGGCTCCACG 1260
Oy 1120 ttcatgtaaaacgctatgcttccacaatttgaatttccacggaatcgctcagcgctccagc 1179
1120 ttcatgtaaaacgctatgcttccacaatttgaatttccacggaatcgctcagcgctccagc 1179
Oy 1261 CTCACCTCCGAATCAATCCATGTAATGCAATACCATGAGACTCGCTCCCTACATGAGTTGCT 1320
1261 ctcacctccgaatcaatccatgtaatgcaataccatgagactcgctccctacatgagttgct 1320
Oy 1180 ctcacctccgaatcaatccatgtaatgcaataccatgagactcgctccctacatgagttgct 1239
1180 ctcacctccgaatcaatccatgtaatgcaataccatgagactcgctccctacatgagttgct 1239
Oy 1321 ACCATGCAACCCCGAGAGATGAATTTAGTTCAGTCCCATCCCTGCTTTGCCGTTAA 1380
1321 accatgcaaccccgagagatgaattttagtTCAGTCCCATCCCTGCTTTGCCGTTAA 1380
Oy 1240 accatgcaaccccgagagatgaattttagtTCAGTCCCATCCCTGCTTTGCCGTTAA 1299
1240 accatgcaaccccgagagatgaattttagtTCAGTCCCATCCCTGCTTTGCCGTTAA 1299
Oy 1381 CCTCATCAAGCTTTTCTGCTGCTTAATCCATACACTGGAATTACCAACTGAGGACATCT 1440
1381 cctcatcaagctTTTCTGCTGCTTAATCCATACACTGGAATTACCAACTGAGGACATCT 1440
Oy 1300 cctcatcaagctTTTCTGCTGCTTAATCCATACACTGGAATTACCAACTGAGGACATCT 1359
1300 cctcatcaagctTTTCTGCTGCTTAATCCATACACTGGAATTACCAACTGAGGACATCT 1359
Oy 1441 ACCCAATACAGAGCTGCGACTGCTCATATGCGTTCATCCATCTGGACGCTACTACTAAG 1500
1441 acccaatacagagctgcgactgctcatatgCGTTCATCCATCTGGACGCTACTACTAAG 1500
Oy 1360 acccaatacagagctgcgactgctcatatgCGTTCATCCATCTGGACGCTACTACTAAG 1419
1360 acccaatacagagctgcgactgctcatatgCGTTCATCCATCTGGACGCTACTACTAAG 1419
Oy 1501 TGGGGAAGAAAGAAA 1516
1501 tggggaaagaaagaaa 1435
Oy 1420 tggggaaagaaagaaa 1435

X

RESULT 2
LOCUS HUMERG2 3166 bp ss-RNA PRI 08-NOV-1994
DEFINITION Human erg2 gene encoding erg2 protein, complete cds.
ACCESSION M17254
KEYWORDS erg 2 protein.
SOURCE Human cell line COLO 320, cDNA to mRNA, clone lambda 12.
ORGANISM Homo sapiens
Eukaryote; Eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 3166)
AUTHORS Rao, V.N., Pappas, T.S. and Reddy, E.S.
TITLE erg, a human ets-related gene on chromosome 21: alternative splicing, polyadenylation, and translation
JOURNAL Science 237 (4815), 635-639 (1987)
MEDLINE 87263429
FEATURES
source
1..3166
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="21q22.3"
257..1645
/gene="ERG"
257..1645
/gene="ERG"

KEYWORDS	c-ets-1 gene; ets gene family; flt-1 gene; integration site.
SOURCE	house mouse.
ORGANISM	Mus musculus
REFERENCE	Eukaryotes: mitochondrial eukaryotes; Metazoa: Chordata; Vertebrata: Eutheria; Rodentia: Sciurognathii; Myomorpha: Muridae; Murine: Mus.
REFERENCE AUTHORS	1 (bases 1 to 1729)
REFERENCE TITLE	Letwin,K. Direct Submission Submitted (08-May-1991) K. Letwin, Samuel Lunenfeld Res Inst at Mount Sinai Hospital, Div of Mol and Developmental Biol, 600 University Avenue, Toronto Ontario M5G 1X5, CANADA
JOURNAL MEDLINE FEATURES	2 (bases 1 to 1729) Ben-David,Y., Giddens,E.B., Letwin,K. and Bernstein,A. Erythroleukemia induction by Friend murine leukemia virus: Insertional activation of a new member of the ets gene family, Flt-1, closely linked to c-ets-1 Genes Dev. 5 (6), 908-918 (1991) 91257578
Location/Qualifiers	1..1729 /organism="Mus musculus" /strain="balb/c" /db_xref="taxon:10090" /germline /cell_type="Friend erythroleukemia" /cell_line="CB7" 1..1729 /gene="Flt-1" /evidence=experimental 1..1729 /gene="Flt-1" 245..1603 /gene="Pli-1" /codon_start=1 /product="retrovirus integration site" /db_xref="PID:950975" /translatio="MDGRIKALSYSDDSLPFSAYGAHLPRADWTAGSPDYGCG PKINLPPOOEWINPYRVAVKRREDDHMGNSRESPDCSSKNCKYKGGEAPMNY NSYDENKGRPPPNMTTNERVITYRPADPTLTQEHVROMLEMAIKELGMIDTSFTQ NMDRELCKMKKEDELRATSAYNTEVLISHLSYRESLLAVNTSHDOSRLNVE DPSTDSVWRGANNNNNNGSLKPSLGGSQTMGNTEBQPDPDYOLIGPSSILANP GSGOILWQFLELLSDSANASCIETEGTNEEFKMTBPDEVARRMGERKSKPMNYDK LSRLRYYYDNKIMTKVGKRAYKFDFHGIALOPHPETSMYKKYPDSISMPSVYR AHQOKVNVPFHSPSSMPVTSSFFGASQYWTSPACIYIPNPVRPHNTHVPSHLS Y"
SE COUNT	466 a 484 C 436 g 343 t
IGIN	
Query Match	26.4% Score 382; DB 29; Length 1729;
Best Local Similarity	67.9%; Prcd. No. 0.00e+00;
Matches	894; Conservative 0; Mismatches 404; Indels 18; Gaps 11;
Db	245 ATGCAGGGACTATTAAAGAGCCTGTCTGTGTGATGAGATCACTCCCTTTTGAT 304 Oy 63 atggagaagcactatgaaggacattatcaagtggtgagtgaagaccagcttcgtttgag 122 Db 305 TCACGATACGAGAGCGGACGCCCATCTCCCCAAGCAGATATGACTGCTTCGGGGAGTCT 364 Oy 123 tgtgtctacgtagtcg-c--cccaacctgcgaagaagaatgacagccttccttcag 179 Db 365 GACTCGGGGACCCCACAAATCAACCCTCGCCACGCGACGACAGAGTGG--ATCA-AC 421 Oy 180 gaatatggccaacaactcaaaagatgagcccgcgtctccccagcagagactggtatcacag 239 Db 422 CAGCCATGAGAGTCAATATGTCACGAGGAGTAGTGCAC--AC-ANG--AATGATCCAG 475 Oy 240 ccccgccagcagtgatcatcatgaagtgtagtggttaaccaccaaacagtgtaaigtggtcaagg 299 Db 476 GAGTCTCCGGTGGACTGCAGTGTCAACAANAATTAACAAGCTGTGTGGCGGAGCGAACCC 535 Oy 300 aattcaactgatgaactgcgttgtgcaaaaagagagaaatgtqtaagcgttcacaaagat 359

RESULT	7	2916 bp	DNA	PAT	05-MAR-1997
LOCUS	A36461				
DEFINITION	Sequence 2 from Patent WO933549.				
D	356	AACCCATGAACATTAATAGCTACATGATGAGAGAGACGGCCCCCTCTCTCCCAACATG	593		
Q	360	gttggagaaactcttggaaactctacatctggaagagacatacttcgcctcctc---aaatag	416		
D	596	ACCCACAACGAAAGGAGAGCTATTGTGCTCCGACAGCCCACTGTGGAGACGAGACAC	655		
Q	417	acaacaaactgaaacgaagagcttaatctgtccagcagatccctcgtctatgagucacagacctt	476		
D	656	GTTCACAGTGTGGCTGGAGTGGGCTCTATAAGAGAAATACGATTTGATGGAGATTCACATTCC	715		
Q	477	gtacgagcaatgtgcgtggagctggcgatgaagagatattgtcttcagagacgtgagacatctg	536		
D	716	TTCTTCGGAACATGATGATGCGACAGAAATTGTGTAATAATAGACAAGAGACTTCCTCCCA	775		
Q	537	tgtctccagaacacatctgattggaagaagcttctgttaaaatgacaaagaatgactctccagaga	596		
D	776	GCCACTCTCGGCTCAACACAGAGAGTCTGTGTGGCCAGCTCAGTTATACCTCAGGAGAAAT	835		
Q	597	ctcaagcccgagctataaagcagataatccctcgtcacaacctaacaactaccctcagagagaga	656		
D	836	TCA-CTGCTGGCTATAACACACMACCTCCATACAGACAGTCTCTCAG-ACGTAATGTGA	893		
Q	657	gagagcaactcttatcttccaaatacatagatttaaccagagcaagcagaaugaatuaa	716		
D	894	AGGAAGACCTCTTCTTATGACTCTCTCAG-GAGAGAGACATGGAACAATATATGACTCT	952		
Q	717	acaagggccaagatttaaccttaagacaaagcgaagagatcagcgtggaacagatcagacat	776		
D	953	GGCCCTCAACAAATGCTCTCTCTCTGGAGATGACAGACCACTGGGCAAGAACATGTAGACAG	1011		
Q	777	cccaactcagtaaaaaagctaacccaacccatctatcttcaaacgtctgccaaacacagagaccag	836		
D	1013	CGGCCCCAGCCAGATCCTTATCAGATCCTGGGGCCAAACGACGCGCTTAGCAAAACCT	1071		
Q	837	cgctctcagtagatcccttaacagatctcttgcagccgacugacgcgtcttgcuaatccca	896		
D	1073	GGGAGTGGGAGATCCAGCGTGGGCACTTCTCTGGAATACTGTGCCAGCGCCGAC	1133		
Q	897	gggagttggcagatacagctatctgtcagcttccactctgtagctcttcttcgagacgtctcaac	956		
D	1133	GCCAGCTGTATCACCTGGAGAGGAGCCAAACGGGGAGTTCAAAATGACGGACCTGTAGAG	1199		
Q	957	tccaactctcatcacctctggaggggacacaaatggggaggtctcaagcttgacagacctgtgaa	1011		
D	1193	GTGGCCAGCGCCTGGGAGAGCGGAGAGCAGACGACCAACATGAATTATGACAACTGACG	1255		
Q	1017	gtggctcgcgcttggggagagagagaaagacaaaccttaacatagaaactatgacaaactcagc	1077		
D	1253	CGGGGCTCCGCATCTACTATGACAAAATAATTATGACCAAACTGGATGGCAAAAGCTAT	1311		
Q	1077	cgtagcaactcctcctactatgaaaaaaatatctatgacttaagcttcaatggtttaaacgtctt	1133		
D	1313	GCTTACAAGTTTGACTTCATGGCATTGCCAGAGGCTCCGACGACACATCCAAACAGAGACA	1377		
Q	1137	gcttcaaaatttgtcttccaagagatcgtctcagggcccttcagctctaacctctcaagatca	1199		
D	1373	TCCATGTACAAGTATCCCTCTGATATCTCTAATCTCTCAATGCTTCTTACATGCGCATCAACAG	1433		
Q	1197	tccaagtacaaatacccaatcagaactccctctacatgattctcttaacatgcuaacccccag	1255		
D	1433	AAGGAGACTTTGTACCGGTCCACCCATACCTTCATGCTGTCACCTCTCTCCACACTCTTT	1499		
Q	1257	aagatgaaacttttgtagctccccaatccccctcgtcttgcctgaatccatccatccagcttttt	1311		
D	1493	GGAGCAGCATACACAATCTGGAGCTCCGCCCACTGCTGGAGTCTTATCCAAACCCGAC	1548		
Q	1317	gtctgcaccttaatccataactgaaattcaacaaactcgaggagcatcttaacccaattaccag	1372		


```
gene /clone_11b="cDNA in lambda gt10"
100..1461
/gene="X1f11"
CDS 100..1461
/gene="X1f11"
/codon_start=1
/product="Xf11 protein"
/db_xref="pid:9505487"
/db_xref="SWISS-PROT:P41157"
/translation="MDGTRKALSVSDQSLFDSAYGASSHLSADMTASANDYQ
PKHINPPODMINQPMKRVNIREYHNSRESVPDCKSIRKSLIGSEGNMTY
TYMDENGPDPMTNERNRVIPADPALMSODHROWLEIKYGLVEIDCSLFON
IDGKELKMSKEDFLRSIVNTEVLHSLNLRDSSSLGYNTOAHDOSRLAKE
DPXYAVRSKSGNMSSPYTKSPMGSTQNNKRSQDPDYQILGPISSLAN
PESGQIQLMQFLLELSDSSNASCITWETNEEFKMTDDEVARKWGEKSPKMYD
KLSRLRYDKSINTKVKGRYAKPDEHQAOLQHPDTSYKYPSEFSYMSY
HSHQKVFVESHSPMSMPTVSSGFFGATSPLYWNSPANLYPMPNRPDTHVQHLG
GFY"
misc_feature 931..1185
/gene="X1f11"
/note="ets-domain"
polya_signal 3449..3455
BASE COUNT 1382 a 836 c 871 g 1314 t
ORIGIN
```

```
Query Match 24.5% Score 355; DB 21; Length 4403;
Best Local Similarity 68.2% Pred. No. 3.15e-295;
Matches 749; Conservative 0; Mismatches 340; Indels 9; Gaps 7;

Db 319 AATGATCCAGGAGTACCGGTGAGCTGAGCATCAACAAATGTAGCAACTATTGGA 378
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 288 aatgggtcaaggaatcactcgtactcagcgtgccaagaaggaagaaatggttagc 347
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 379 GGGTCGGAGGCAATGCAATGACTTACCTACATGATGAAAGAATGGCCCCCTCCC 438
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 348 agttcagaacaatggttggaactatggaagctacatgaagaagacatattccgct 407
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 439 CCAATATATACATCAAAATGAAGAGGGTTATTGTACCTGCAGATCCCGTTGTGGAGC 498
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 408 ccaaatatgacaacacatgaacgaagaatgattgtgcagcagatcctactgattgagc 467
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 499 CAGGACAGTACGAGGAGTGGGTGAGTGAATGAATGAAATGAGTGTGGAAAT 558
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 468 aagaacaatgacagcagcagcagcagcagcagcagcagcagcagcagcagcagc 527
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 559 GACTGCTCCCTCTTCAGATATGATGCAAGAACTGTGTAATATGACAAAGAGAT 618
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 528 gcaatctgtgttccagaaatgattggaagagtggtgtataaattgacaaagattgac 587
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 619 TTCTTGGGAGCAGCATCATATATTAACACAGAGTTTACTTTTCATCTCAACTACTTC 678
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 588 ttccagaagactcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 647
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 679 AGGGATAGTAGTCATCATTTGGTTATACACTCAGGACATCAACAAATATGAGCT 738
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 648 agagaagaagagagccacttatttccaaataacatcagtttaaccagaagaagacaa 707
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 739 TTGACAGCCA-AAGAGATCCCT-CCT-ATGAAGCGTAAGAGATCTGGATGGGGAAT 795
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 708 agaataaacaacaagcagcagcagcagcagcagcagcagcagcagcagcagcagc 767
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 796 TCCATGAGCTCACTGTCACAAAAAGTCTCCCATGGAGGAACCAAAAATGTGAATAA 855
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 768 caca-gccatc--ccactcagtcacaagcacaac--atcattcacaagcagccaaa 824
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 856 ACCGGGATCAGCAAGCTTCTGACCAATCCATACCAATATATAGGGCCCAACAAGCAG 915
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 825 aagaagaagacagc--gtccctcagttagatccttcaatctttagaccgagcagcagc 881
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 916 ACAGTGGCTAAACCAAGGAGTGGCAAAATTCAGCTGTGGAGTGTGGAGTGTGCA 975
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 882 cgtcttgcaaatccagcagcagcagcagcagcagcagcagcagcagcagcagcagc 941
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Db 976 TCTGACAGTTCAAAGCTAGTGTATCATGAGGAGCACTAATGAGAGTTCAAAATG 1035
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 942 tcgagacagcccaactccaactcactcagcagcagcagcagcagcagcagcagcagc 1001
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1036 ACCGACCCAGATGAAGTGAAGCGAGCTTGGGTGAAGAAAGCAAGCTTAATGAAAC 1095
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1002 acagaacccagatgaagatgctcgcgtggggagagagagagagagaaacctaactgac 1061
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1096 TATGACAAATTAAGCCGAGCACTAAGATATCTATTTGATAAAGTAAATGACTTAAAGTC 1155
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1062 tatgacaactcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1121
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1156 CATGGTAAAGCTATGACCCACAAATTTGACTTTCACGGCATGTCAGGACATGAGGCT 1215
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1122 cactggaacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1181
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1216 CATCCAACTGATACACTCATGATGACAAATATCTTGTGAGTTCTTACATGCTTCTTAC 1275
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1182 caccctccagaatcattcattgacaaataccatagacatcccttaccatggttccctac 1241
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1276 CATTCACATCAACAGAGGTTAATTTTGTCCCTTCACACCCATCCTCTATGCGGTAACA 1335
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1242 catgcaacccccagagaagatgaaattgtagctcccatcccccctgcttgccgtaacc 1301
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1336 TCATCAGGCTTCTTGGAGCAACATCTCCTTACTGGAATTCACCAAGTCAAACTTAT 1395
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1302 tcatcagcgttcttgcgtcccttaactcatcagtaattcaccatcaggaagcatctac 1361
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1396 CCAAAACCCAAATGTGCCA 1413
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1362 cccaataccagcgtgcga 1379
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 11
LOCUS AP000022 133746 bp DNA PRI 13-MAY-1998
DEFINITION Homo sapiens genomic DNA, chromosome 21q22.2 (Down Syndrome
region), segment 15/15, complete sequence.
ACCESSION AP000022
NID 93132332
KEYWORDS HTG.
SOURCE Homo sapiens DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 133746)
AUTHORS Hirakawa,M., Yamaguchi,H. and Imai,K.
TITLE Homo sapience 1,296,826bp genomic DNA of 21q22.2 Down Syndrome
region
JOURNAL Published Only in Database (1998) In press
REFERENCE 2 (bases 1 to 133746)
AUTHORS Hirakawa,M., Yamaguchi,H. and Imai,K.
TITLE Direct Submission
JOURNAL Submitted (11-MAY-1998) to the DDBJ/EMBL/GenBank databases. Mike
Hirakawa, Japan Science and Technology Corporation (JST), Advanced
Databases Department, 5-3, Yonban-cho, Chiyoda-ku, Tokyo, Tokyo
102-0028, Japan (E-mail:mikae@tokyo.jst.go.jp, Tel:81-3-5214-8491,
Fax:81-3-5214-8470)
```

FEATURES

source	1..133746
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="21"
	/map="21q22.2"

BASE COUNT 38362 a 29023 c 28295 g 38066 t

ORIGIN

```
Query Match 24.4% Score 353; DB 27; Length 133746;
Best Local Similarity 83.7% Pred. No. 2.47e-293;
Matches 438; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

Db 33592 TTATAGTAGTAAAGCCCAAGTGAAGGATGATGCTGTGGGAGCCATGTTGGGGA 33651
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Cp 1418 ttatagtagtgcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1359
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

BASE COUNT	846 a	698 c	683 g	727 t	ORIGIN
Query Match	23.6%	Score 341;	DB 26;	Length 2954;	
Best Local Similarity	67.9%;	Pred. No. 5,68e-282;			
Matches 873;	Conservative	0;	Mismatches 388;	Indels 24;	Gaps 14
Db	173	ATGCACGGGACTATTAAAGAGCCTCTGCTGGTGTGAGCGACGACCACTGCTTTTGAC	232		
QY	63	atgcgaagcactataaagaagcattatcaagtgtgtgtagtgaaacagtcctctttttg	122		
Db	233	TCACGATACGGAGGGGACGCCCACTCTCCCAAGGCCCAACATGATGCTCTGGGGAGTCT	292		
QY	123	tgtpccctcggatcg-c-cccaccttgcagaagacagaanaagaaagcctcttcacgt	179		
Db	293	GACTACGGGGGACCCCAACAGATCAACCCCTCCACACACAGCAGAGATGG--ATCA-AT	349		
QY	180	gaatatggygcaaacatcaagaatgaagcccgcggtctcccgacgagactgtttatcaag	239		
Db	350	CAGCCAGTAGAGGGTCAACGCTCAACGGGAGATGATGAC--AC-ATG--AATGATCCAG	403		
QY	240	ccccggcgcaagatctaccattaaagtatgtagtgaatcaaccaacacgattaatggtcaag	299		
Db	404	GAGTCTCCGGTGTGACTGACGCGTTAAGCAATGCAAGCAAGCTGTGTGGCGGAGGGAGTCC	463		
QY	300	aattacacttgatgactgcagcgctggcacaagaaggaagaaatggtttagcaatcagacat	359		
Db	464	AACCCATGAGACTCAACACAGCTAATATGGAGCAGAAAGATGGCCCCCTCCCAACATG	523		
QY	360	gttggaatgaactatggaagctacatggaaagaagaactatccgcccce--aatatg	416		
Db	524	ACCAACCAACGAGAGAGATGATCGTCCGCCCAACACCAACATCTGTGGACACAGAGCAT	583		

[illegible]

QY 1014 gaagtgcgcgcgttcggggaaggaagaaacgaacctaacatgaaactatgacaaatc 1073
Db 1178 AGCCGGGCCCCCTCCGTTATTACTATGATATAAAGATTATGACAAAGTGACAGGCCAAGA 1337
QY 1074 agcgtgcgaacttcgcctactactatgacaaaataatatagtactaaagtcatgtgaaacgc 1133
Db 1238 TATGCTTACAAATTGACTTCCACGGCATTCGCCAGCTCTGCGACCAATCCGACGAG 1297
QY 1134 tatgcctacaataattgattcctcaggaatcgtcgaagccctcagcctcaccctcagaa 1193
Db 1298 TCGTCCATGACAGTACGCTTCCATCATCTCCATACATGCGCTTCCCAACATGCCACAG 1357
QY 1194 tcttcctatgacaataatccatccatgacctccctcctcctcctcctcctcctcctcctc 1253
Db 1358 CAGAAAGTGAAGTCTTGTCCCTCCCATCATCTCCATGCGCTCTGCTCCAGCTTC 1417
QY 1254 cagaagatgaaactttgtagtccctccatccctccctcctcctcctcctcctcctcctc 1313
Db 1418 TTTGGAGCCCGCATCACAATCTGGA 1442
QY 1314 ttctgtccccctaatcactactga 1338

RESULT 13
LOCUS S72865 588 bp mRNA PRI 24-JAN-1995
DEFINITION EMS. . EMS-erg-EMS-erg fusion protein type 9e [human, SK-PN-LI cell
line, mRNA Partial Mutant, 3 genes, 588 nt].
ACCESSION S72865
NID 9633777
KEYWORDS human SK-PN-LI cell line.
SOURCE Homo sapiens
ORGANISM Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 588)
REFERENCE
AUTHORS Giovannini, M., Biegel, J.A., Serra, M., Wang, J.Y., Wei, Y.H.,
Nycum, L., Emanuel, B.S., and Evans, G.A.
EMS-erg and EMS-Fli1 fusion transcripts in Ewing's sarcoma and
primitive neuroectodermal tumors with variant translocations
J. Clin. Invest. 94 (2), 489-496 (1994)
JOURNAL
MEDLINE 94314948
REMARK Genbank staff at the National Library of Medicine created this
entry [NCBI g1db15q154042] from the original journal article.
This sequence comes from Fig. 2A.
FEATURES
source
1..588
location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..588
/partial
/note="EMS-erg fusion protein type 9e"
1..588
gene
/partial
/gene="EMS"
1..588
/partial
/gene="erg"
1..588
CDS
/partial
/gene="EMS-erg"
/note="EMS-erg fusion protein type 9e; This sequence comes
from Fig. 2A"
/codon_start=1
/db_xref="PID:9633778"
/translation="PTSTPQTSTSYQAPSOYSQSSSTGGQSSGQIQMQFLLELLS
DSSNSSCITMEGNGEKMTPDEVARRMGERSKENMNYDKLSRLRYYYDNIMTK
VHGKRAYKDFHGIADQHPHPRESSLYKPPSDLPYMGSKYHAHPOKMFVADHPAL
PVTSSSEFAAPNYMNSPTGTYNTRLPSTSHPSH"

BASE COUNT 145 a 210 c 129 g 104 t
ORIGIN
Query Match 23.3%, Score 337; DB 26; Length 588;

Best Local Similarity 83.6%; Pred. No. 3,466-278;
Matches 419; Conservative 0; Mismatches 82; Indels 0; Gaps 0;
Db 88 AGTGGCAGATCCAGCTTTGGCAGTTCTCTCGAGACTCTGTGAGACAGTCCAACTCC 147
QY 900 agtggcagataagctatagctatggcagttcctactcgtgagcttcctgtcgacagctccaaatcc 959
Db 148 AGCTGCATCAGCTTGGGAGGACCAAGGAGGAGTTCAAGATGACGAGTCCGACGAGTG 207
QY 960 aactgcataccctcgtggagggacaatctggaggttcaagatgacagacccctgtgaugtg 1019
Db 208 GCCCGGCGCTGGGAGAGCGGAGAACCAACCAATGATACATGATAGCTACAGCCGC 267
QY 1020 gctcggcgttgaggagagagaaagcaaacctcaacatgaacttgaacaactcagcgct 1079
Db 268 GCCCTCGTTATCATGATGACAGAACATCATGACCAAGGTCATGGAGGCTACGCGC 327
QY 1080 gcaactcgtactactatgacaaataatataatgacaaagtccaatgtaaaagctctgc 1139
Db 328 TACAGTTGAGATTCAGAGGAGATCGCCAGGCCCTCCAGCCCGCCGAGATCATCT 387
QY 1140 tacaatttgatttcccaaggaatcgtcagggccctccagcctcaccctcagaaatcatcc 1199
Db 388 CTGTACAGTACCCCTCAGACCTCCCTGTACATGGCTCTATACGCCACCAAGCAAG 447
QY 1200 atgtacaataatcacatcagaactccctccatcagttcctacatgacacaccccgagag 1259
Db 448 ATGAGCTTTGTGGCGGCCGCCCTCCAGCCCTCCGCTGATCTTCAGTTTGTGCT 507
QY 1260 atgaacttctgagctccctccctccctcctcctcctcctcctcctcctcctcctcctc 1319
Db 508 GCCCAACCCATCATGATGATTCACCAAGCTGGGGGTATATACCAACATGAGCTCCGC 567
QY 1320 gccctcaatcacatcaggaattcaccaactcagggagcatctacccaataaccagggcgca 1379
Db 568 ACCAGCCATATGCTTCTCAT 588
QY 1380 gctgcataatgcttcccat 1400

RESULT 14
LOCUS HUMFLI1A 1932 bp mRNA PRI 31-DEC-1994
DEFINITION Human Fli-1 mRNA, complete cds for two alternate splicings.
ACCESSION M93255
NID 9182659
KEYWORDS Fli-1; alternative splicing; ets oncogene family; transcription
factor.
SOURCE Homo sapiens (tissue library: lambda g111 HBL 2T) blood cDNA to
mRNA.
ORGANISM Homo sapiens
Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1932)
REFERENCE
AUTHORS Hromas, R., May, W., Denny, C., Raskind, W., Moore, J., Maki, R.A.,
Beck, E., and Klemsz, M.J.
Human Fli-1 localizes to chromosome 11Q24 and has an aberrant
transcript in neuroepithelioma
JOURNAL Blochim. Biophys. Acta 1172 (1-2), 155-158 (1993)
MEDLINE 93176799
FEATURES
source
1..1932
location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="HEL"
/cell_type="erythro leukemia"
/tissue_type="blood"
/tissue_lib="lambda g111 HBL 2T"
/map="11q23-25"
join(1..401,576..1932)
/gene="Fli-1"
/note="alternate"
1..1932
/note="alternate"

```
gene
      join(1..401,576..1932)
      /gene="FLI-1"
      join(174..401,576..1700)
CDS
      /gene="FLI-1"
      /note="alternate"
      /codon_start=1
      /db_xref="PID:g182660"
      /translation="MDGTLKALSVSDSDSLFSDAYGAANHLPRADMTASGSPDYQ
      PHKINLPPOEENINOPRVKRVKDYDMHNSSESPVDCSVKSKLVGGSGSNPNNTSYDEKN
      SYDEKNGPPPNMTNERRIVPADATLVTOEHVROMLEWAIKEYSLMEIDTFEQMDKELE
      MDKELEKMKNKEDLRATTLTNTLYVLSHLSYLESSILANTSHTDQSRRLVKED
      PSTDSVRGAMNGNSGLNKSPPGAGQITISKNEQDPDPQIIGPISRLANPESGOIQW
      SGOIQWQFLELLSDSANSCITMEGTNGEFTKMTDEVARRMGORKSKNNMYDKL
      SRAIRYKDKNIMTKVHKRYAKKFDHFGINAOLQPHTESSMKYRPSDISYMSYAHQOKVNF
      HQOKVNFVPHPSMPTVSSFFGASQYMTSTGCIYFNPVPRHPTNHPVSHLSGYX
      "
      543..1700
      /gene="FLI-1"
      /note="alternate"
      /codon_start=1
      /product="FLI-1"
      /db_xref="PID:g182661"
      /translation="MEGSLAGEFARRESVDCSVKSKLVGGSGSNPNNTSYDEKN
      GPPPNMTNERRIVPADATLVTOEHVROMLEWAIKEYSLMEIDTFEQMDKELE
      KMKNEDELRATTLTNTLYVLSHLSYLESSILANTSHTDQSRRLVKEDPSYDVR
      RGAMNGNSGLNKSPPGAGQITISKNEQDPDPQIIGPISRLANPESGOIQW
      OFLELLSDSANSCITMEGTNGEFTKMTDEVARRMGORKSKNNMYDKL
      YDKNIMTKVHKRYAKKFDHFGINAOLQPHTESSMKYRPSDISYMSYAHQOKVNF
      VPHPSMPTVSSFFGASQYMTSTGCIYFNPVPRHPTNHPVSHLSGYX"

BASE COUNT      487 a      554 c      519 g      372 t
ORIGIN
chromosome 11q23-25.

Query Match      22.5%      Score 325      DB 26      Length 1932:
Best Local Similarity 68.3%      Pred. No. 7.63e-267:
Matches 719:      Conservative 0:      Mismatches 322:      Indels 12:      Gaps 7:
```

```

      826 cagaagaccagcgtccctcagttagatcccttaccatctcttgaccgaccgacgctc 885
      112 TAGCCAAACCCGTGGAAGGGGGAGATCCAGCTGTGGCAATTCTCTGTGAGTGCCTCG 1221
      886 ttgcaatccaggaggtggcagatacagctatggcagttccctactgagcttcgtcgg 945
      1222 ACAGGCCCAAGCCGACGCTGTATACCTCTGGAGGGGAGCCAGGGGAGTTCAAATGACGG 1281
      946 acagctccaactccaactgcgatacctctgaggaggacaaatgggaggttcaagatgacag 1005
      1282 ACCCGATGAGGTGGCCAGCGCGCTGGGGCAGCGGAAAGCAAGCCCAATGAAATTACG 1341
      1006 accctgagaaagtgcgtcggttgaggagaggaagaagaacactaacatgaaactatg 1065
      1342 ACAAGCTAGCGCGGCGCTCGTATTACTATGATTAATAAACAATTATGACCAAGTGCACG 1401
      1066 acaactcagccgtgcactcgtcactctatgacaaaataatattatgactaaagttcatg 1125
      1402 GCAAAAGATATGCTTACAAATTTGACTTCACAGGCATGCCCAGGCTGTGCAGCCACATC 1461
      1126 gtaaacgctatagccatacctaattgatttccaagagatcgtccagccctccagctcaac 1185
      1462 CGACCGAGTCTGCATGTACAGTACAGTACCTTCTGACATCTCTACATGCTTCTTACCATG 1521
      1186 ctccgaatcatcatcatgatacaataaccatacagactccctccatcatgattcctacatg 1245
      1522 CCCACGACAGAAAGTGAACCTTGTCCCTCCCAATCCATCCCTCATGCTGCTCATCTCT 1581
      1246 cacaccccccaagaatgaacttctgattgattcccatccctctgcttggccgttaacctat 1305
      1582 CCAGCTCTTGGAGCGCATCATCAATACTGGA 1614
      1306 ccagctttttgtctccctcaatccatctcgtga 1338

RESULT 15
LOCUS      CCRNAFLI      3490 bp      mRNA      VRT      07-JUL-1998
DEFINITION      Ccrnafi ccrnafi mRNA for transcription factor FLI, clone 10.1.
ACCESSION      Y14773
KEYWORDS      g3269302
NID      93269302
KEYWORDS      FLI oncprotein; transcription factor.
SOURCE      quail.
ORGANISM      Coturnix coturnix
Eukaryota; Metazoa; Chordata; Vertebrata; Archosauria; Aves;
Neognathae; Galliformes; Phasianidae; Phasianinae; Coturnix.
REFERENCE      1 (bases 1 to 3490)
Mager,A.M., Grapin-Botton,A., Iadjaali,K., Meyer,D., Wolff,C.M.,
Stiegler,P., Bonnin,M.A. and Remy,P.
The avian fl gene is specifically expressed during embryogenesis
in a subset of neural crest cells giving rise to mesenchyme
Int. J. Dev. Biol. 42 (4), 561-572 (1996)
JOURNAL      98358003
MEDLINE      2 (bases 1 to 3490)
REFERENCE      Stiegler,P.
AUTHORS      Direct Submission
TITLE      Submitted (09-SEP-1997) P. Stiegler, Centre National de la
Recherche, Scientifique, U.P.R. 9005 du C.N.R.S., 15 Rue Rene
Descharre, 67084 Strasbourg Cedex, FRANCE
COMMENT      Related sequence Y14774.
FEATURES
      source
      1..3490
      /organism="Coturnix coturnix"
      /db_xref="taxon:9091"
      /chromosome="1"
      /dev_stage="4-day quail embryo"
      /clone_1lb="pcDNA1"
      /clone="10.1"
      /map="g31-33"
      215..1513
      /gene="FLI1"
      215..1513
      /gene="FLI1"
      1102 ATACAGAGCAAGCGCCCGCCAGCATCCTATCAGATCCTGGGCCGACGACGATGCGCC 1161
```

```

/function="transcription factor"
/codon_start=1
/product="FLI transcription factor"
/db_xref="PID:g3308365"
/db_xref="PID:g3308365"
/db_xref="PID:g3308365"
/translaltion="MOST IDEAL SVYSDOSLEFSTYTGATHTLPRADMTASGNDYGO
PHTKNLPPEPOEWIMQCVRVNRVAREYEDHNHNGSSESVDCSNVAKCRVLGATGTEENPMS
YSTIADMRKNGPPEMTINERVARYPADPITLOQEHNRQALAEYGLMIDITTFPO
MMDKEKCEKMKKODELFTLYTLYEVLITSLRSSSLATYANPSTSESSSLARK
EGPVAGVAGNOGKTEOORROPOYLLITSLRSSLATPSSGQIOLMOLLELLSDSN
ASCTIWTGNTGEPRMTDPEDVARNRMEERKSPMAMTDKSLRLRYTYRKNMTYTBHKK
RYAKEDPHGLAQLQHPHLESSMYTSDLSIMSTYRAHQKQVNFVPHHSPMPTYS
SSFGAASPVTYSAGSIVYVWPNVPHANHVSPLHSGTY"

```

/codon_start=1

/product="FLI transcription factor"

/db_xref="PID:g3269303"

```

/translation="MDGTIKEAL

```

YSTYMDENKONPPRMTTNEBRYVPADPTPLTNGEHRQWLEAKIKEYGMLBIDTTFQ
 MNDREELKKNKIDTLELJYVTEVLESLRASSLYNTNPTSTESLSRLAK
 EGPVAVGIONKIDTEQRQPOPIYILPTSRSLNPSGCOIQMLCELESDSSNK
 ASCITWETNGEEMDPDVARAMEKSRPMYNDKSLARYDYKMTVYHKK
 RYACTWETGIAQLOPHPLESSNYTPEDSLSTMSYTAHQOQKVFDPHFSNAPVTS
 SIFKAECPWTSPAGISIVNPNVPHRPAHVSPLHSGY

a	770 c	752 g	942 t
---	-------	-------	-------

22.18; Score 320; DB 21; Length 3490;

ity 79.28; Pred. No. 4.03e-262;

inservative 0; Mismatches 114; Indels 0; Gaps 0;

.....

BASE COUNT	1026 a	770 c	752 g	942 t
ORIGIN				

ORIGIN

Query Match	22.18;	Score 320;	DB 21;	Length 3490;
-------------	--------	------------	--------	--------------

Best Local Similarity 79.28; Pred. No. 4.03e-262;

```
Matches 434; Conservative 0; Mismatches 114; Indels 0; Gaps 0;
```

910 AACGAAACACACGCGCCCAACACAGATTCATATCAAAATCTGGGGCCCAACAGTAGTCG 969
 ||||| | | | | | ||||| | | | | | ||||| | | |
 824 AACGGAAGACCAAGCTCCTCAATGATACCTTATCAGATCTCTTGAACGACCAAGCAGCGCG 883

QY 824 aaacgaagaccagcgtcctcagttagatccttatcagattcttgacgaccagcagccg 883

Db 970 TCTGCCAATCCTGGGAGTGGGCAGATACAACCTGTGGCAGTTCCTCCTCGAGTTGCTGTC 10299

QY 884 tcttgcaatccaggagtggcagatcacgctatgagcttcctactgagcttctgtc 943

Db 1030 GGACAGTTCCAATGCCAGCTGTATCACATGGGAAGGACCATGGGAATTCAGATGAC 1089

QY 944 ggacagctccaactccaactgcatcaactggagggcacaaatggggagtccaagatgac 100

D5 1090 AGACCCAGATGAAGTGGCACGGCTGGGGAGAACGCAAAAGCAAGCCCAACATGATTA 1149

1004 agaccctgagtcggcctcggcgttggggaggaagaacaaacctaacatgaaacta 1063

DD 1150 1GACAAAC 1GAGCCGAGCCCI 1CGAIAIC 1ACIA 1GAIAGAAACAI 1AIGALCAAAGI 1GCA 1205

VY 1004 cycccccccllcagllcylcyrcacccclylccalcacccacgacccaaaaaacalllaalgcrladaagyllcca .1123

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

.....

[illegible][illegible]

1304	1305	1306	1307	1308	1309	1310	1311	1312	1313	1314	1315	1316	1317	1318	1319	1320	1321	1322	1323	1324	1325	1326	1327	1328	1329	1330	1331	1332	1333	1334	1335	1336	1337	1338	1339	1340	1341	1342	1343	1344	1345	1346	1347	1348	1349	1350	1351	1352	1353	1354	1355	1356	1357	1358	1359	1360	1361	1362	1363	1364	1365	1366	1367	1368	1369	1370	1371	1372	1373	1374	1375	1376	1377	1378	1379	1380	1381	1382	1383	1384	1385	1386	1387	1388	1389	1390	1391	1392	1393	1394	1395	1396	1397	1398	1399	1400
1304	1305	1306	1307	1308	1309	1310	1311	1312	1313	1314	1315	1316	1317	1318	1319	1320	1321	1322	1323	1324	1325	1326	1327	1328	1329	1330	1331	1332	1333	1334	1335	1336	1337	1338	1339	1340	1341	1342	1343	1344	1345	1346	1347	1348	1349	1350	1351	1352	1353	1354	1355	1356	1357	1358	1359	1360	1361	1362	1363	1364	1365	1366	1367	1368	1369	1370	1371	1372	1373	1374	1375	1376	1377	1378	1379	1380	1381	1382	1383	1384	1385	1386	1387	1388	1389	1390	1391	1392	1393	1394	1395	1396	1397	1398	1399	1400

nh 1450 2A2CCCCA 1457

OV 1364 caatacca 1371

Search completed: Sat Apr 10 10:21:10 1999
Job time : 3786 secs.

Job time : 3786 secs.

[illegible]

RESULT	2		
ID	Q14319	PRELIMINARY;	PRT; 385 AA.
	Q14319;		
DT	01-NOV-1996	(TREMBLREL, 01, CREATED)	
DT	01-NOV-1996	(TREMBLREL, 01, LAST SEQUENCE UPDATE)	
DT	01-JUN-1998	(TREMBLREL, 06, LAST ANNOTATION UPDATE)	
DE	FLI-1.		
GN	FLI-1.		
OS	HOMO SAPIENS (HUMAN).		
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;		
OC	EUTHERIA; PRIMATES.		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=BLOOD;		
RX	MEDLINE; 93176799.		
RA	HROMAS R., MAY W., DENNY C., RASKIND W., MOORE J., MAKI R.A., BECK E.		
RA	KLEMSZ M.J.		
RL	BIOCHIM. BIOPHYS. ACTA 1172:155-158(1993).		
DR	EMBL; M93255; G18261; -		
DR	PROSITE; PS00345; ETS_DOMAIN1; 1.		
DR	PROSITE; PS00346; ETS_DOMAIN2; 1.		
PRAM	PF00178; Ets.		
SEQ	SEQUENCE 385 AA; 43437 MW; 0820C243 CRC32;		

Query Match	40.48;	Score 1401;	DB 4;	Length 385;
Best Local Similarity	56.48;	Pred. No. 9.42e-281;		
Matches	229;	Mismatches 73;	Indels 35;	Gaps 21

[illegible]

RESULT	3	PRELIMINARY:	PRT:	254	AA.
ID	016199				
AC	Q16199;				
DT	01-NOV-1996	(TREMBL:REL. 01, CREATED)			
DT	01-NOV-1996	(TREMBL:REL. 01, LAST SEQUENCE UPDATE)			
DT	01-JUN-1998	(TREMBL:REL. 06, LAST ANNOTATION UPDATE)			
DE	EMS-ERG FUSION PROTEIN TYPE 1E (FRAGMENT).				
GN	EMS-ERG.				
OS	HOMO SAPIENS (HUMAN).				
CC	EUAROTIA; METAFOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;				
OC	EUTHERIA; PRIMATES.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE: 94314948.				
RA	GIOVANNINI M., BIEGEL J.A., SERRA M., WANG J.Y., WEI Y.H., NYICUM L.,				
RA	EMANUEL B.S., EVANS G.A.;				
RL	J. CLIN. INVEST. 94:489-496(1994).				
DR	EMBL: S72621; G633774; -.				
DR	PROSITE: PS00345; ETS_DOMAIN_1; 1.				
DR	PROSITE: PS00346; ETS_DOMAIN_2; 1.				
DR	PFAM: PF00178; ETS.				
FT	NON_TER	1			
SO	SEQUENCE	254	AA;	256580CD	CRC32;
Query Match	36.5%	Score 1267;	DB 4;	Length 254;	
Best Local Similarity	77.7%;	Pred. No. 5.03e-250;			
Matches 176;	Conservative	20;	Mismatches 28;	Indels 3;	Gaps 3

Query Match	36.5%	Score 1267	DB 4	Length 254
Best Local Similarity	77.7%	Pred. No. 5	OB 20	
Matches	178	Conservative	20	Mismatches 28; Indels 3; Gaps 3
Db	27	QONTPEPPRRSANTGCHGHPPOSKAAPSESTVPKTEDQRPOLDYQIIGPSSRLANP	86	
QY	247	RPDPPEQARRSASMTSHSHPT-QSKAQOPSSSTYPKTEDQRPOLDYQIIGPSSRLANP	305	
Db	87	GSGQIQLOQFLLELLSDSSNSCITWEGTNEEFYKTPQDEYARHWGRKSKPNNNYKLS	146	
QY	306	GSGQIQLOQFLLELLSDSSNSNCITWEGTNEEFMTQPDDEYARHWGRKSKPNNNYKLS	365	
Db	147	RALRYDYDKNMTKTVGKRYAKFDFGIAQALQPHRPESLSYKYP-SDLPYMGSYAHNP	205	
QY	366	RALRYDYDKNMTKTVKHPRESS-MYKYPQSDLPYMGSSYGKRYAKFDFGIAQALQPHANP	424	
Db	206	QKMNVAHPALPYVTSSFFAANPNPNYSTGIGITNTPLPTSHMESH	254	
QY	425	QKMNVAHPALPYVTSSFFAANPNPNYSTGIGITNTPLPTSHMESH	473	

RESULT	4			
ID	Q16203	PRELIMINARY;	PRT;	196 AA.
AC	Q16203;			
DT	01-NOV-1996	(TREMBLREL, 01, CREATED)		
DT	01-NOV-1996	(TREMBLREL, 01, LAST SEQUENCE UPDATE)		
DT	01-JUN-1998	(TREMBLREL, 06, LAST ANNOTATION UPDATE)		
DE	EMS..	EMS-ERG-EMS-ERG FUSION PROTEIN TYPE 9E (FRAGMENT).		
GN	EMS-ERG.			
OS	HOMO SAPIENS (HUMAN).			
OC	EUDAROTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;			
CC	EUTHERIA; PRIMATES.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE; 94314948.			
RA	GIOVANNINI M., BIEGL J.A., SERRA M., WANG J.Y., WEI Y.H., NYCOM L.,			
RA	EMANUEL B.S., EVANS G.A.;			
RL	J. CLIN. INVEST. 94:489-496(1994).			
DR	EMBL; S72865; G633778; -			
DR	PROSITE; PS00345; ETS_DOMAIN_1; 1.			
DR	PROSITE; PS00346; ETS_DOMAIN_2; 1.			
DR	PRFAM; PF00178; Ets.			
FT	NON_TER	1		
SQ	SEQUENCE	196 AA; 22268 MW; FBDC632E CRC32;		
Query March	26.4%;	Score 915;	DB 4;	Length 196;
Best Local Similarity	73.3%;	Pred. No. 5,52e+10;		
Matches	129;	Conservative 30;	Indels 2;	Gaps 2

[illegible]

[illegible]

Query Match	15.78;	Score 543;	DB 11;	Length 340;
Best Local Similarity	73.78;	Pred. No. 2.35e-87;		
Matches	70;	Conservative	13;	Mismatches 10;
				Indels 2;
				Gaps 2

Dd 137 GLPSAVOK -SSGQIOLQFLLELLADANGCJAMEEGHSEFFLTDPDEVARMRGRKS 195
 : : : : :
 Oy 286 GPTSRRLANPESGOQLWQFLELLSDSNSNCITWEETNEEFKMTDPDEVARMGRKS 355
 Db 196 KPNMNYDKLSRALRYDYDKNTMSKYHGCRYA -YRF 229
 : : : : :
 Oy 356 KPNMNYDKLSRALRYDYDKNTMTKVHPRESSMYFY 390

```

RESULT      9
ID          091744      PRELIMINARY;      PRT;      268 AA.
AC          091744;
DT          01-NOV-1996 (TREMBLREL. 01, CREATED)
DT          01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT          01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE          XEI-A DNA-BINDING PROTEIN (FRAGMENT).
GN          XEI-A.
OS          XENOPUS LAEVIS (AFRICAN CLAWED FROG).
OC          EUARHOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AMPHIBIA; ANURA
RN          [1]
RP          SEQUENCE FROM N.A.
RC          TISSUE:TOTAL OVARY FROM SINGLE FROG;
RA          SALVATI F., MORABITO, MEREDINDO, CARREVALLI;
RL          SUBMITTED (JUL-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
DR          EMBL; X65167; G517228; -
DR          PROSITE; PS00345; ETS_DOMAIN_1; 1.
DR          PROSITE; PS00346; ETS_DOMAIN_2; 1.
DR          PFAM; PF00178; ETS.
KW          DNA-BINDING.
FT          NON_TER      1
SQ          SEQUENCE      268 AA;      30805 MW;      39A1C89D CRC32;

```

```
Query Match      13.3%    Score 460;   DB 13     Length 268;  
Best Local Similarity 44.4%;    Pctd. No. 1,84e-6;  
Matches 64; Conservative 3; Mismatches 45; Indels 2  
  
Db      110 HDSFSEDDYPP-AMPSHKSQCFKRYVDRALNNDCKRVIPAAAGACTGSGPQLMQFL 168  
       : : | :| :||:::| :| :||||: : ||| :|||
```

[illegible]

RESULT	10	
ID	062804	PRELIMINARY;
NC	062804	PRT; 336 AA

DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE TRANSCRIPTION FACTOR GABP ALPHA SUBUNIT (FRAGMENT).
OS OVIS ARIES (SHEEP).
OC EURKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA.
OC EUTHERIA; ARTIODACTYLA.
RN [1]
RP SEQUENCE FROM N.A.
RA EZASHI T., EALY A.D., OSTROWSKI M.C., ROBERTS R.M.:
RL SUBMITTED (APR-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AF057717; G3046690; -
DR PROSITE; PS00345; ETS_DOMAIN_1; 1.
DR PROSITE; PS00346; ETS_DOMAIN_2; 1.
FT NON_TER 1
SQ SEQUENCE 336 AA; 38033 MW; 8CFEE91F CRC32;

Query Match	13.1%;	Score 454;	DB 6;	Length 336;
Best Local Similarity	46.7%;	Pred. NO. 3.51e-68;		
Matches	63;	Conservative	32;	Mismatches 34;
			Indels	6;
			Gaps	5;

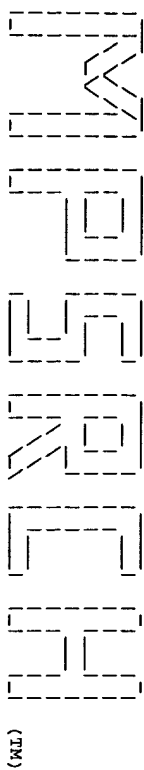
```

Db      163  PTAIAIN-SSAKAKAKYO-RAPRTISGDRSP--GNRTGN--NGOIOLOMPELLLELTDKDA 217
Oy      266  PTOASATPPSSSTVPKTEDEORQDPPYOLIGPSSSRLANEGSGOIOLOMPELLLELSSSN 325
Db      218  RDCISWVGDSEEPFLNOPELVAOKMGORKNKPNTNYESLSPALRYYYDGMICKVOG--KR 276
Oy      326  SNCITWEGTNEEPKMTDDEVAARRMGEKSKPNNMYDKLSRALRYYYDKNIMTKVHPES 385
Db      277  FYVKFVCDLTKTLIGY 291
Oy      386  SMTKIPSDLPYMSST 400

```

ID	026645	PRELIMINARY;	PRT:	559 AA.
AC	026645;			
DT	01-NOV-1996	(TREMBLEL. 01. CREATED)		
DT	01-NOV-1996	(TREMBLEL. 01. LAST SEQUENCE UPDATE)		
DT	01-JUN-1998	(TREMBLEL. 06. LAST ANNOTATION UPDATE)		
DE	ETS HOMOLOGUE.			
OS	STRONG/LOECENTROTUS PURPURATUS (PURPLE SEA URCHIN).			
OC	EUKARYOTA; METAZOA; ECHINODERMATA; ECHINOZOA; ECHINOIDEA.			
OC	EUCHINOIDEA.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=EMBRYONIC;			
RA	RAO S. K., CHILDS G.;			
RL	SUBMITTED (JUN-1993) TO EMBL/GENBANK/DBJ DATA BANKS.			
DR	EMBL: L19541: G310662; _			
DR	PROSITE; PS00345; ETS_DOMAIN_1; 1.			
DR	PROSITE; PS00346; ETS_DOMAIN_2; 1.			
DR	PFAM; PF00178; ETS.			
SO	SEQUENCE 559 AA; 62052 MW; 0F896532 CRC32;			

Query Match	13.0%;	Score 449;	DB 5;	length 559;
Best Local Similarity	58.0%;	Pred. No. 4.09e-67;		
Matches	58;	Conservative	18;	Mismatches 22; Indels 2; Gaps 2



Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MSPrch_p protein - protein database search, using Smith-Waterman algorithm
Run on: Wed Apr 7 09:34:36 1999; Maspar time 26.24 Seconds
856,003 Million cell updates/sec

Title: >US-08-878-177-2
Description: (1-451) from US08878177.ppt
Perfect Score: 3286
Sequence: 1 MASTIKELSVSEDSQSLFE.....IYPNRLPRAHMPSHLGTY 451

Scoring table: PAM 150
Gap 11

Searched: 165420 segs, 49795644 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: spiremb16
1:sp.archaea 2:sp.bacteria 3:sp.fungi 4:sp.human
5:sp.invertebrate 6:sp.mammal 7:sp.mhc 8:sp.organelle
9:sp.phage 10:sp.plant 11:sp.protein 12:sp.unclassified
13:sp.vertebrate 14:sp.virus

Statistics: Mean 47.907; Variance 81.767; scale 0.586

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	2989	91.0	478 13	Q90837	ERG PROTEIN.	0.00e+00
2	1973	60.7	385 4	Q14319	FLI-1.	0.00e+00
3	1600	48.0	254 4	Q16199	EMS-ERG FUSION PROTEIN	0.00e+00
4	1248	38.0	196 4	Q16202	EMS.. EMS-ERG-EMS-ER	1.17e-263
5	730	22.2	188 5	Q22355	COSMID T08H4.	3.82e-138
6	668	20.3	238 4	Q99581	FEV PROTEIN.	2.08e-123
7	664	20.3	340 11	Q70132	ETS DOMAIN TRANSCRIPT	2.08e-123
8	664	20.2	208 5	Q18579	SIMILARITY TO ETS DOMA	1.84e-122
9	650	19.8	336 6	Q62804	TRANSCRIPTION FACTOR G	3.78e-119
10	494	15.0	268 13	Q91744	XEL-A DNA-BINDING PROT	1.14e-82
11	483	14.7	426 6	Q62803	TRANSCRIPTION FACTOR E	3.91e-80
12	483	14.7	559 5	Q26645	ETS HOMOLOGUE.	8.54e-73
13	451	13.7	393 5	Q61805	C. ELEFANS LIN-1 (GB:U	8.54e-73
14	451	13.7	441 5	Q27378	LIN-1.	2.76e-70
15	440	13.4	494 13	Q57586	ETS-DOMAIN TRANSCRIPT	2.30e-67
16	427	13.0	471 13	Q57586	TRANSCRIPTION FACTOR G	1.31e-53
17	366	11.1	300 11	Q70273	EHF (EHF).	1.28e-49
18	348	10.6	533 4	Q15725	NERF-1A.	1.28e-49
19	348	10.6	533 4	Q15724	NERF-1B.	1.28e-49
20	348	10.6	581 4	Q15723	ETS TRANSCRIPTION FACT	1.28e-49

Result ID	Score	Query Match	Length	ID	Description	Pred. No.
21	349	10.6	663 4	Q99607	MYELOID ELF-1 LIKE FAC	7.70e-50
22	349	10.6	663 4	Q60435	ELF-1 RELATED PROTEIN.	7.70e-50
23	344	10.5	481 5	Q17057	C24A1.2 PROTEIN.	9.77e-49
24	336	10.2	76 4	Q16198	EMS.. FLII (FRAGMENT	5.63e-47
25	330	10.0	81 14	Q64998	(TSL1 MUTANT) V-ETS (1.17e-45
26	313	9.5	242 5	Q01521	SIMILARITY TO THE DNA-	5.95e-42
27	304	9.3	348 4	Q99718	EPITHELIAL-SPECIFIC TR	5.29e-40
28	304	9.3	371 11	Q35275	ETS TRANSCRIPTION FACT	5.29e-40
29	304	9.3	371 4	P78545	ESE-1B.	5.29e-40
30	307	9.3	532 5	Q19685	SIMILAR TO DNA BINDING	1.19e-40
31	288	8.8	70 4	Q16324	TL5.. ERG (TRANSDUCA	1.46e-36
32	269	8.2	314 5	Q01519	SIMILARITY TO THE DNA-	1.59e-32
33	262	8.0	42 6	Q62767	ERYTHROBLASTOSIS VIRUS	4.73e-31
34	264	8.0	507 5	Q93320	HYPOTHETICAL PROTEIN C	1.80e-31
35	250	7.6	180 5	Q44138	Q50A2.4 PROTEIN.	1.52e-28
36	245	7.5	146 13	Q91921	XREF1ALPHA (FRAGMENT).	1.65e-27
37	243	7.4	33 5	Q25773	ETS-1 PROTEIN (FRAGEN	4.29e-27
38	228	6.9	33 5	Q16997	ETS-2 PROTEIN (FRAGEN	5.09e-24
39	227	6.9	33 5	Q25573	ETS-1 PROTEIN (FRAGEN	8.12e-24
40	227	6.9	33 5	Q25574	ETS-2 PROTEIN (FRAGEN	8.12e-24
41	227	6.9	98 11	Q55199	ETS TRANSCRIPTION FACT	8.12e-24
42	225	6.8	33 5	Q25127	ETS-1 PROTEIN (FRAGEN	2.07e-23
43	225	6.8	33 5	Q25066	ETS-1 PROTEIN (FRAGEN	2.07e-23
44	225	6.8	33 5	Q23822	ETS PROTEIN (FRAGMENT	2.07e-23
45	224	6.8	33 5	Q25159	ETS-1 PROTEIN (FRAGEN	3.30e-23

ALIGNMENTS

Result ID	Score	Query Match	Length	ID	Description	Pred. No.
1	2989	91.0%	478 13	Q90837	ERG PROTEIN.	0.00e+00
AC	Q90837	Best Local Similarity 94.4%;	Pred. No. 0.00e+00;			
AD	Q90837	Matches 451;	Conservative 0;	Mismatches 0;	Indels 27;	Gaps 2;
DT	01-NOV-1996 (TREMURREL. 01, CREATED)					
DT	01-NOV-1996 (TREMURREL. 01, LAST SEQUENCE UPDATE)					
DT	01-JUN-1998 (TREMURREL. 06, LAST ANNOTATION UPDATE)					
DE	ERG PROTEIN.					
GN	ERG.					
OS	GALLUS GALUS (CHICKEN).					
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;					
OC	GALLIFORMES.					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	TISSUE-SPLEEN;					
RX	MEDLINE: 95329425.					
RA	DIORDAIN P., DEWITTE F., DESBIENS X., STEHELIN D.,					
RA	DUTERQUE-COQUILLAUD M.;					
RL	MECH. DEV. 50:17-28(1995).					
DR	EMBL: X77159; G790440; -.					
DR	PROSITE: PS00345; ETS_DOMAIN_1; 1.					
DR	PROSITE: PS00346; ETS_DOMAIN_2; 1.					
DR	PRAM; PF00178; Ets.					
SQ	SEQUENCE 478 AA; 53913 MW; CFEAB9C2 CRC32;					
Query Match	Score 2989; DB 13; Length 478;					
Best Local Similarity 94.4%;	Pred. No. 0.00e+00;					
Matches 451;	Conservative 0;	Mismatches 0;	Indels 27;	Gaps 2;		
Db	1 MASTIKELSVSEDSQSLFECAVSPHLAKTEMASSSESGTSSKSPRPQDWSQP 60					
Qy	1 MASTIKELSVSEDSQSLFECAVSPHLAKTEMASSSESGTSSKSPRPQDWSQP 60					
Db	61 PARTIMECPNPNVNSRNSPDCSAKAGKAMVSSSDNNGMNTGSTMEKHIPPNTT 120					
Qy	61 PARTIMECPNPNVNSRNSPDCSAKAGKAMVSSSDNNGMNTGSTMEKHIPPNTT 120					
Db	121 NERVIVPADPTLSTHVRHOMLEWAVEYGLPVDVILLFQNDKELCKTKTDFORT 180					
Qy	121 NERVIVPADPTLSTHVRHOMLEWAVEYGLPVDVILLFQNDKELCKTKTDFORT 180					
Db	121 NERVIVPADPTLSTHVRHOMLEWAVEYGLPVDVILLFQNDKELCKTKTDFORT 180					
Qy	121 NERVIVPADPTLSTHVRHOMLEWAVEYGLPVDVILLFQNDKELCKTKTDFORT 180					
Db	181 PSYNADILSLHLYRLRPLPHLTSDDVDKALQNSPRLMHARNTGATFFPNTSVPEA 240					
Qy	181 PSYNADILSLHLYRLRPLPHLTSDDVDKALQNSPRLMHARNTGATFFPNTSVPEA 240					
Db	181 PSYNADILSLHLYRLRPLPHLTSDDVDKALQNSPRLMHARNTGATFFPNTSVPEA 240					
Qy	181 PSYNADILSLHLYRLRPLPHLTSDDVDKALQNSPRLMHARNTGATFFPNTSVPEA 240					

DB 241 TORITRBDLPEQARRSAMTSHSPOTOSKATQBPSSSTVPRKEDORPOLDPYQIIGPTSS 300
QY 214 TORITRBDLPEQARRSAMTSHSPOTOSKATQBPSSSTVPRKEDORPOLDPYQIIGPTSS 273
DB 301 RLANGSGOIQLOMFLLELSDSSNSCITWEGTNGEFKMTDPDVARWGRKSKPMNN 360
QY 274 RLANGSGOIQLOMFLLELSDSSNSCITWEGTNGEFKMTDPDVARWGRKSKPMNN 333
DB 361 YDKLSRALRYYYDKNIMTKVHGKRAYAKFDFHGAIALQPHPESSMYKYPSPDLPYMSY 420
QY 334 YDKLSRALRYYYDKNIMTKVHGKRAYAKFDFHGAIALQPHPESSMYKYPSPDLPYMSY 393
DB 421 HAHPOKMFVAPHPALPVYSSFFAAPNPYNSPTGIGYPTRLPAHMPSHLGTYY 478
QY 394 HAHPOKMFVAPHPALPVYSSFFAAPNPYNSPTGIGYPTRLPAHMPSHLGTYY 451

RESULT 2
ID Q14319 PRELIMINARY: PRT: 385 AA.
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE FLI-1.
GN FLI-1.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAFOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
CC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE: BLOOD;
RX MEDLINE: 9316799.
RA HROMAS R., MAY W., DENNY C., RASKIND W., MOORE J., MAKI R.A., BECK E.,
RA KLEMSZ M.J.;
RL BIOCHIM. BIOPHYS. ACTA 1172:155-158(1993).
DR PROSITE; PS00345; ETS_DOMAIN_1; 1.
DR PROSITE; PS00346; ETS_DOMAIN_2; 1.
DR PFM; PF00178; Ets.
SQ SEQUENCE 385 AA; 43437 MW; 0820C243 CRC32;

Query Match 60.0%; Score 1973; DB 4; Length 385;
Best Local Similarity 67.8%; Pred. No. 0.00e+00;
Matches 257; Conservative 68; Mismatches 46; Indels 8; Gaps 6;

DB 10 ARESVDCSVSKSLVGGESNPNNYSYMDKNGPPNNNTTNERIYVPADATLVTO 69
QY 78 SRNSDDCSVAKGKAMVSSSDVNGMNGSYMEKH-IPPNMTTERKIYVADPTLMST 136
DB 70 EHVROMLEMAKEYSLMEIDTSFFQNDGKELCKNNKEDFLRATTLVNTVLLSHLSTYL 129
QY 137 DHVROMLEMAKEYSLMEIDTSFFQNDGKELCKNNKEDFLRATTLVNTVLLSHLSTYL 196
DB 130 E-SLLAY-NTSHTDQSSRLSVKEDPSYDVRGAMGNMNSGLKSPPLGCAOTISN 187
QY 197 ERGAFFIRPNISVYDEAQRITRBDLPEQARRSAMTSHSPOTOSKATQBPSS-TVPKT 255
DB 188 TEOROPDPYOLIGPTSSRLANGSGOIQLOMFLLELSDSANSICITWEGTNGEFKMTD 247
QY 256 EDORQOLPQYOLIGPTSSRLANGSGOIQLOMFLLELSDSANSICITWEGTNGEFKMTD 315
DB 248 PDEVARRGOKRSKPMNMYDKLSRALRYYYDKNIMTKVHGKRAYAKFDFHGAIALQPH 307
QY 316 PDEVARRGOKRSKPMNMYDKLSRALRYYYDKNIMTKVHGKRAYAKFDFHGAIALQPH 375
DB 308 TESSMYKPPSDISYMPSTHAHQKVFVPPHSSMPVTSSTFFGAAGQWTS-TGGIYFN 366
QY 376 PESSMYKPPSDISYMPSTHAHQKVFVPPHSSMPVTSSTFFGAAGQWTS-TGGIYFN 435
DB 367 PNVPRHPTHVPSHLGSYY 385
QY 436 T---RLPAHMPSHLGTYY 451

RESULT 3
ID Q16199 PRELIMINARY: PRT: 254 AA.
AC Q16199;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE EMS-ERG FUSION PROTEIN TYPE 1E (FRAGMENT).
GN EMS-ERG.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAFOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
CC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 94314948.
RA GIOVANNINI M., BIEGEL J.A., SERRA M., WANG J.Y., WEI Y.H., NYCUM L.,
RA EMANUEL B.S., EVANS G.A.;
RL J. CLIN. INVEST. 94:489-496(1994).
DR EMBL; S72621; G633774; -;
DR PROSITE; PS00345; ETS_DOMAIN_1; 1.
DR PROSITE; PS00346; ETS_DOMAIN_2; 1.
DR PFM; PF00178; Ets.
FT NON-TER 1
SQ SEQUENCE 254 AA; 28577 MW; 596580CD CRC32;

Query Match 48.7%; Score 1600; DB 4; Length 254;
Best Local Similarity 93.4%; Pred. No. 0.00e+00;
Matches 213; Conservative 12; Mismatches 2; Indels 1; Gaps 1;
DB 27 QONLPYEPFRSANTGCHHPFPOSKAOPSPSTYPKTDORPOLDPYQIIGPTSSRLANP 86
QY 220 RBDLPYEPFRSANTGCHHPFPOSKAOPSPSTYPKTDORPOLDPYQIIGPTSSRLANP 278
DB 87 GSGOIQLOMFLLELSDSSNSCITWEGTNGEFKMTDPDVARWGRKSKPMNMYDKLS 146
QY 279 GSGOIQLOMFLLELSDSSNSCITWEGTNGEFKMTDPDVARWGRKSKPMNMYDKLS 338
DB 147 RALRYYYDKNIMTKVHGKRAYAKFDFHGAIALQPHPESSMYKYPSPDLPYMSYHAHQ 206
QY 339 RALRYYYDKNIMTKVHGKRAYAKFDFHGAIALQPHPESSMYKYPSPDLPYMSYHAHQ 398
DB 207 KMFVAPHPALPVYSSFFAAPNPYNSPTGIGYPTRLPAHMPSH 254
QY 399 KMFVAPHPALPVYSSFFAAPNPYNSPTGIGYPTRLPAHMPSH 446

RESULT 4
ID Q16203 PRELIMINARY: PRT: 196 AA.
AC Q16203;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE EMS-ERG FUSION PROTEIN TYPE 9E (FRAGMENT).
GN EMS-ERG.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAFOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
CC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 94314948.
RA GIOVANNINI M., BIEGEL J.A., SERRA M., WANG J.Y., WEI Y.H., NYCUM L.,
RA EMANUEL B.S., EVANS G.A.;
RL J. CLIN. INVEST. 94:489-496(1994).
DR EMBL; S72621; G633778; -;
DR PROSITE; PS00345; ETS_DOMAIN_1; 1.
DR PROSITE; PS00346; ETS_DOMAIN_2; 1.
DR PFM; PF00178; Ets.
FT NON-TER 1
SQ SEQUENCE 196 AA; 22268 MW; F8CD632E CRC32;

Query Match 38.0%; Score 1248; DB 4; Length 196;
Best Local Similarity 93.7%; Pred. No. 1.17e-263;
Matches 164; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

```

Db 22 SSSYGGSSGGIOIOMQFLELLSSSSNCITWGTNGEEMTDDBDVARWGKSKPN 81
Oy 272 SSRLANPSSGGIOIQMQLFLELLSSSSNCITWGTNGEEMTDDBDVARWGKSKPN 331
Db 82 MNVKLSPALRYYYDKNIMTKVHGKRVAYKDFGIAQALOPHPPESSLYKYPEDLYMG 141
Oy 332 MNVOKLSALALYYDDKNIMTKVHGKRVAYKDFGIAQALOPHPPESSMYKYPEDLYMS 391
Db 142 SYHAHPQKMNVAADHPALPYTSSSFFAAPNPYNNSPGTCGIPYNTRLPTSPMPSH 196
Oy 392 SYHAHPQKMNVAADHPALPYTSSSFFAAPNPYNNSPGTCGIPYNTRLPTSPMPSH 446

RESULT 5
ID 022355 PRELIMINARY: PRT: 188 AA.
AC 022355;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
FE COSMID 108H4.
FE T08H4.3.
FE CAENOSHADITIS ELEGANS.
OC EUAROTATA: METAQOA; ACOELOMATES; NEMATODA; SECCERNMENTEA; RHADITIDA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
EX MEDLINE: 94150718.
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M., BONFIELD J.,
RA BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A., CRAXTON M.,
RA DEAR S., DU Z., DURBIN R., FAVELLO A., FULLON L., GARDNER A., GREEN P.,
RA HARKINS T., HILLIER L., JETER M., JOHNSON L., JONES M., KERSHAM J.,
RA KIRSTEN J., LAISTER N., LATREILLE P., LIGHTNING J., LLOYD C.,
RA KICURERAY A., KOPPIKORE B., O'CALLAGHAN M., PARSONS J., PERCY C.,
RA RIKKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R., SMALDON N., SMITH A.,
RA SONNHAMMER E., STAEDEN K., SUSTON J., THIERRY-MING J., THOMAS K.,
RA VANDIN M., VAUGHAN K., WATERSON R., WATSON A., WEINSTOCK L.,
RA WILKINSON-SPEOAT J., WOHLDMAN P.;
RL NATURE 368:32-38 (1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA STELLYES L.;
RL SUBMITTED (NOV-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA WATERSTON R.;
RA SUBMITTED (OCT-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
EMBL: U39470; G1041871; -.
RA PROSITE: PS00345; ETS_DOMAIN_1; 1.
DR PROSITE: PS00346; ETS_DOMAIN_2; 1.
DR PFAM: PF00178; Ets.
SQ SEQUENCE 188 AA; 21345 MW; C98FD1EC CRC32;

Query Match 22.2%; Score 730; DB 5; Length 188;
Best Local Similarity 82.8%; Pred. No.3.82e-138;
Matches 96; Conservative 8; Mismatches 11; Indels 1; Gaps 1.

Db 5 DPYOILGPTSKNLHSGSGGOTQMQFLELLSDKRYSEVITWEGTQGEFKLVDEDEYARK 64
Oy 263 DPYOILGPTSSRLANPSSGGIOIQMQLFLELLSDSSNSNCITWGTNGEEMTDDBDVARR 322
Db 65 WGERSKRNMMYDKNRSLRATRYYYKKNIMAKVHGKRYAKKPFQOIAALOP-PRAS 119
Oy 323 WGERSKRNMMYDKNRSLRATRYYYKKNIMTKVHGKRYAKKPFQOIAALOPHPPESS 378

RESULT 6
ID 099581 PRELIMINARY: PRT: 238 AA.
AC 099581;
DT 01-MAY-1997 (TREMBLREL. 03, CREATED)
DT 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)

```

Query Match	20.3%	Score 668;	DB 11;	Length 340;
Best Local Similarity	77.3%	Pred. No. 2,08e-123;		
Matches	85;	Conservative 15;	Mismatches 9;	Indels 1;
Db	34	GPLSPAYOK-GSGQIQIQLMQLFLELLADLRANAGCIAMEGHEGFEKLTDPDEVARRMGERKS	92	
Qy	269	GPTSSRLANPSSGQIQIQLMQLFLELLSDSSNSNCITGWGTNGEFKMTDPDEVARRMGERKS	328	
Db	93	KPNMNYDKLSRALRYDYDKNIMTSVHGKRYAYRFPDFOGLAACOPPPAHA	142	
Qy	329	KPNMNYDKLSRALRYDYDKNIMTVHGKRYAYRKFDFGIAQALPDPHPES	378	
RESULT	7	PRELIMINARY;	PRT:	340 AA.
ID	070132			
AC	070132			
DT	01-AUG-1998	(TREMBLREL. 07, CREATED)		
DI	01-AUG-1998	(TREMBLREL. 07, LAST SEQUENCE UPDATE)		
DE	01-AUG-1998	(TREMBLREL. 07, LAST ANNOTATION UPDATE)		
DS	ETS DOMAIN TRANSCRIPTION FACTOR PET-1.			
OS	RATUUS NORVEGICUS (RAT).			
OC	EUDAROTIA; METAPODA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;			
CC	EUTHERIA; RODENTIA.			
RP	[1]			
RN	SEQUENCE FROM N.A.			
RC	TISSUE-PHEOCHROMOCYTOMA;			
RX	MEDLINE; 98127904.			
RA	FYODOROV D., NELSON T., DENNERIS E.;			
RL	J. NEUROBIOL. 34:151-163(1998).			
DR	EMBL: U91679; G3033419; -			
DR	PROSITE; PS00345; ETS_DOMAIN_1; 1.			
DR	PROSITE; PS00346; ETS_DOMAIN_2; 1.			
SO	SEQUENCE 340 AA; 35439 MW; B3214B28 CRC32;			
Query Match	20.3% <td>Score 668;<td>DB 11;<td>Length 340;</td></td></td>	Score 668; <td>DB 11;<td>Length 340;</td></td>	DB 11; <td>Length 340;</td>	Length 340;
Best Local Similarity	77.3% <td>Pred. No. 2,08e-123;</td> <td></td> <td></td>	Pred. No. 2,08e-123;		
Matches	85; <td>Conservative 15;<td>Mismatches 9;<td>Indels 1;</td></td></td>	Conservative 15; <td>Mismatches 9;<td>Indels 1;</td></td>	Mismatches 9; <td>Indels 1;</td>	Indels 1;
Db	137	GPLSPAYOK-GSGQIQIQLMQLFLELLADLRANAGCIAMEGHEGFEKLTDPDEVARRMGERKS	195	
Qy	269	GPTSSRLANPSSGQIQIQLMQLFLELLSDSSNSNCITGWGTNGEFKMTDPDEVARRMGERKS	328	
Db	136	KPNMNYDKLSRALRYDYDKNIMTSVHGKRYAYRFPDFOGLAACOPPPAHA	245	
Qy	329	KPNMNYDKLSRALRYDYDKNIMTVHGKRYAYRKFDFGIAQALPDPHPES	378	
RESULT	8	PRELIMINARY;	PRT:	208 AA.
ID	Q18579			
AC	Q18579			
DT	01-NOV-1996	(TREMBLREL. 01, CREATED)		
DI	01-NOV-1996	(TREMBLREL. 01, LAST SEQUENCE UPDATE)		
DE	01-JUN-1998	(TREMBLREL. 06, LAST ANNOTATION UPDATE)		
DS	SIMILARITY TO ETS DOMAINS.			
GN	C42D8.4.			
OS	CANORHABDITIS ELEGANS.			
OC	EDAROTIA; METAPODA; ACCELOMATES; NEMATODA; SECERNENTEA; RHABDITIDA.			

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2.
RA MEDLINE: 94150718.
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M., BONFIELD J.,
RA BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A., CRAXTON M.,
RA DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L., GARDNER A., GREEN P.,
RA HAKINS T., HILLIER L., JIER M., JOHNSTON L., JONES M., KERSHAM J.,
RA KIRSTEN J., LAISTER N., LATREILLE P., LIGHTNING J., LLOYD C.,
RA MCORRAV A., MORTMORE B., O'CALLAGHAN M., PARSONS J., PERCY C.,
RA RIFFEN L., ROOPRA A., SANDERS D., SHOWNKEEN R., SMALDON N., SMITH A.,
RA SONNHAMMER E., STADEN R., SULSTON J., THIERRY-MIEG J., THOMAS K.,
RA VAUDIN M., VAUGHAN K., WATERSTON R., WATSON A., WEINSTOCK L.,
RA WILKINSON-SPROAT J., WOHLDMAN P.,
RA NATURE 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2.
RA HALLSWORTH K.
RA SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2.
RA WATERSTON R.
RL EMBL: U56966; G1293845; -.
DR PROSITE: PS00345; ETS_DOMAIN_1; 1.
DR PROSITE: PS00346; ETS_DOMAIN_2; 1.
DR PFAM: PF00178; Ets.
SQ SEQUENCE 208 AA; 22664 MW; 3492097D CRC32;

Query Match 20.28; Score 664; DB 5; Length 208;
Best Local Similarity 79.48; Pctd. No. 1.84e-122;
Matches 81; Conservative 14; Mismatches 7; Indels 0; Gaps 0;

Db 21 GPMALASATGTGQIQLOMFLLELLADAVAHGICWNGEFLVDPDEVARWGERKS 80
Qy 269 GPTSSSLANPGSQIQLOMFLLELLSDSSNCKITWEGTNGEFLVDPDEVARWGERKS 328
Qy 329 KPMNMYDKLSRALRYYYDKNIMTKVQGRYAYKDFGGLAQA 122
Qy 329 KPMNMYDKLSRALRYYYDKNIMTKVQGRYAYKDFGGLAQA 370

RESULT 9
ID 062804; PRELIMINARY; PRT: 336 AA.
AC 062804;
DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
DE TRANSCRIPTION FACTOR GABP ALPHA SUBUNIT (FRAGMENT).
OS OVIS ARIES (SHEEP).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; ARTIODACTYLA.
RN [1]
RP SEQUENCE FROM N.A.
RA EZASHI T., EALY A.D., OSTROMSKI M.C., ROBERTS R.M.;
RL SUBMITTED (APR-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: AF057717; G3046690; -.
DR PROSITE: PS00345; ETS_DOMAIN_1; 1.
DR PROSITE: PS00346; ETS_DOMAIN_2; 1.
FT NON_TER 1
SQ SEQUENCE 336 AA; 38033 MW; 8CFEE91F CRC32;

Query Match 19.88; Score 650; DB 6; Length 336;
Best Local Similarity 41.68; Pctd. No. 3.78e-119;
Matches 101; Conservative 61; Mismatches 61; Indels 20; Gaps 15;

Db 58 EEBELGIPYDPIQWSTDQVLAHVWVWVWKEFSMTDILHTLL-NISGELCSLSOEDEFQRY 116
Qy 121 NBRVYVADPDLMTSTDVHOMLEMAVKEVGLPDVILLFQNDIGKELCKMTKDDF-QRL 179
Qy 117 -PR--GEILMSHLELLR-K---YVLASQF--QOMNE-IYTI-DQP-VQIIPAS-VQSAIP 163
```

```

Qy 180 TPEYNADILLSHLHYLRREGATFIFFPNYSYPEATQRIITTPDLPYEQARRSAMTSHSP 239
Db 164 TAIKAIN-SKAKAYO-RAPRISGEDRSSP-GNRTGN--NGQIQLOMFLLELLTDKAR 218
Qy 240 TQSKATQPSSTVPKTEDDQRPQDIPYQIIGPTSSRLANPGSQIQLOMFLLELLSSSS 299
Db 219 DCSWVGDEGEFRLNPELYAQMGQRKPKPYNTEKLSRALRYYYDGDMICKVQGRV 278
Qy 300 NCTIWTGTEGFEKMTDPDEVARWGERKSPPNNYDKLSRALRYYYDKNIMTKVHGKRYA 359
Db 279 YKF 281
Qy 360 YKF 362

RESULT 10
ID 091744; PRELIMINARY; PRT: 268 AA.
AC 091744;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
DE XEL-A DNA-BINDING PROTEIN (FRAGMENT).
GN XEL-A.
OS XENOPUS LAEVIS (AFRICAN CLAWED FROG).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AMPHIBIA; ANURA.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-TOTAL OVARY FROM SINGLE FROG;
RA SALVATI F., MORABITO, MERENDINO, CARNEVALI;
RL SUBMITTED (JUL-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X65167; G517228; -.
DR PROSITE: PS00345; ETS_DOMAIN_1; 1.
DR PROSITE: PS00346; ETS_DOMAIN_2; 1.
DR PFAM: PF00178; Ets.
KW DNA-BINDING.
FT NON_TER 1
SQ SEQUENCE 268 AA; 30805 MW; 39A1C89D CRC32;

Query Match 15.08; Score 494; DB 13; Length 268;
Best Local Similarity 49.68; Pctd. No. 1.14e-82;
Matches 66; Conservative 28; Mismatches 38; Indels 1; Gaps 1;

Db 110 HDSFSEDYPP-AMPSHKSKGTFKDYVRDAELINKDKPIVPAALAGYTGSGPIQLOMFL 168
Qy 230 RSAMTSHSHPTQSKATQPSSTVPKTEDDQRPQDIPYQIIGPTSSRLANPGSQIQLOMFL 289
Db 169 LELLTQKSCQSFISWTDGDEWFKLSPDEVARWGERKSPPNNYDKLSRALRYYYDKNI 228
Qy 290 LELLSDSSNSNCKITWEGTNGEFLVDPDEVARWGERKSPPNNYDKLSRALRYYYDKNI 349
Db 229 IHKTAGKRYVYRF 241
Qy 350 MKTVHGKRYAYKF 362

RESULT 11
ID 062803; PRELIMINARY; PRT: 426 AA.
AC 062803;
DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
DE TRANSCRIPTION FACTOR ETS-2 (FRAGMENT).
OS OVIS ARIES (SHEEP).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; ARTIODACTYLA.
RN [1]
RP SEQUENCE FROM N.A.
RA EZASHI T., EALY A.D., OSTROMSKI M.C., ROBERTS R.M.;
RL SUBMITTED (APR-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: AF057716; G3046688; -.
DR PROSITE: PS00345; ETS_DOMAIN_1; 1.
DR PROSITE: PS00346; ETS_DOMAIN_2; 1.
```



```

/translation="MASTIKALSVSEDSLEFACVSPHIAKTEMTASSSEYQOT
SKMSRVPQODMLSQPARVITKMECNNOVNGSNSSNDVAGKRVSSSDNVGM
NYGSMEKRIEPPMNTNERVIVADPTLSTEDHAKLEWAKVKGKIVSDILLE
ONIDKELCKMTKDFORLSTPBNADILSHLHYRETPILTSDDYDKALONSPRL
MHARTGATRTFPNTSYPRATORITTRPDLPYOARBSAATSSHDPKOSANQPS
STVPTEDQRFOLDYQILGPTSSRLANPGSQIDLMQFLELSDSSNSNITTEGT
NGEFTMDPDEVARRWGERKSKPMNMYDKLSHALYIYDKNMTYVHGKRYIKRDFH
GIAQALQHPPESSMYKYPSDLPYMSYHAHPQKNFVAHPHPALPVSSSFEAPNP
YWSPTGGIYPTRLPAAHMPHLGTYT"
BASE COUNT      458 a      392 c      327 g      339 t
ORIGIN

```

```

Query Match      99.2%; Score 1516; DB 21; Length 1516;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 1516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 GAATCCGGAGCAAGTAATTTATTAGCAATTTAGCGATCAATTAATCTTGATCACA 60
    1 gaaatccggagcaagaaataattatattagcaaatattagcgatcaaatcttgatcaca 60
    61 TTATGGCAAGCACTATTAAAGAGCAATTTACGTGTGATGAGCAAGCAAGCTCTGTTTG 120
    61 ttaaggcaagcaactattaaagagcaatttacgtgtgatgagcaagcaagctcttgattg 120
    61 ttaaggcaagcaactattaaagagcaatattagcgatcaaatcttgatcaca 60
    121 AGTGTGCTACAGGATGCGCCACCTTGCAAAAGACAGAAATGACAGCCTCTCTCCAGTG 180
    121 agtgtgctacaggatgccccaccttgcaaaagacagaatgacagcctctctccagtg 180
    121 agtgtgctacaggatgccccaccttgcaaaagacagaatgacagcctctctccagtg 180
    181 AATATGGCAACATCAAGATAGCCCGCGCTTCCCGACGAGACTGGTTATCAGAC 240
    181 aatattggcaaacatcaagatagcccgcgcttcccgacgagactggttatcagac 240
    181 aatattggcaaacatcaagatagcccgcgcttcccgacgagactggttatcagac 240
    241 CCCCAGGAGAGTTACATTAAGATGAGTGTAAACCAACAGGTTAAAGGGTCAAGA 300
    241 cccagaggagagttacattaaagatgagtgtaaccacaacaggttaaggggtcaaga 300
    241 cccagaggagagttacattaaagatgagtgtaaccacaacaggttaaggggtcaaga 300
    301 ATTACCTGATGATCAGCGGTGGCAAAAGAGGAAATGTTAGCAGTTGACAGACATG 360
    301 attacctgatgatcagcggtggcaaaagaggaaatgttagcagttgacagacatg 360
    301 attacctgatgatcagcggtggcaaaagaggaaatgttagcagttgacagacatg 360
    361 TTGGGATGAATCTATGCAAGCTACATGGAAGAAAGCAATATCCGCTCCAAATATGCAA 420
    361 ttgggatgaatctatgcaagctacatggaagaaagcaaatatccgctccaaatatgcaa 420
    361 ttgggatgaatctatgcaagctacatggaagaaagcaaatatccgctccaaatatgcaa 420
    421 CCAATGACGAAGAGTATTGTGCACAGATCTTACTTATGGAGACAGACACATATAC 480
    421 ccaatgacgaagagtattgtgcacagatcttacttattggagacagacacatatac 480
    421 ccaatgacgaagagtattgtgcacagatcttacttattggagacagacacatatac 480
    481 GGCAGTGGCTGAGTGGGAGTGAAGAGTATGTTCTTCAGAGCTGAGACATCTTGTGT 540
    481 ggcagtggctgagtgggagtgaaagtattgttcttcagagctgagacatcttgtgt 540
    481 ggcagtggctgagtgggagtgaaagtattgttcttcagagctgagacatcttgtgt 540
    541 TCCAGAAATGATGAGGAAAGAGTGTGTAATATGACCAAAAGATGACTTCCAGAGACTCA 600
    541 tccagaaatgatgaggaagagtggttaaatatgacccaaaagatgacttccagagactca 600
    541 tccagaaatgatgaggaagagtggttaaatatgacccaaaagatgacttccagagactca 600
    601 CGCGAGGCTTAACGCGAGATATCTCTGTACACACTACCTACCTACGAGAGACTCTC 660
    601 cgcgaggcttaacgcgagatattctctgtacacactacctacactcagagagactctc 660
    601 cgcgaggcttaacgcgagatattctctgtacacactacctacactcagagagactctc 660
    661 TTCACATTTGACTTTCAGATGATGTTGATAAGGCTTACAAAACCTCCACGGTAAATGC 720
    661 ttccacatcttgactttcagatgatggttataaggttataagacctccacggtaaatgc 720
    661 ttccacatcttgactttcagatgatggttataagacctccacggtaaatgc 720
    721 ATGCTAGAAAACACAGAGAGCCATTTTATTTTCCAAATACATCAGTTTACCAGAG 780
    721 atgctagaaaacacagagagccatttttccaaatatacagtttaccagag 780
    721 atgctagaaaacacagagagccatttttccaaatatacagtttaccagag 780
    781 CAACGCAAAATATACACAAGGCGACATTTACTTATGAGCAAGGACAGAGATCAGCT 840
    781 caacgcaaaatatacacaagcgacatttacttataagcaagcgagagatcagcgt 840
    781 caacgcaaaatatacacaagcgacatttacttataagcaagcgagagatcagcgt 840
    841 GGAGAGTCAAGCCATCCCATGTCACAAAGATACCAACATCATCTTCAACAGTGC 900

```

```

OY 841 ggcagagtcagcagcattccactcagtcacaaagctacccaacacacacacacagtc 900
    901 CCAAAAACAGAACAGCAGCGTCTCTCAGTTAGATCTTATAGATTTCTTGAGCCAGCACA 960
    901 ccaaaaacagaaacagcagcgctcctcagttagatcattacagttcttgagccagcagca 960
    961 GCCGCTCTTGCAAAATCCAGGAGTGGGAGATACAGCTATGCGAGTTCCTACTGAGCTTC 1020
    961 gccgctcttgcaaaatccaggagtcggagatcacagctatgagttcttactcagtcctc 1020
    1021 TGTGGACAGCTCCCACTCACTGATTCACCTGAGGAGGCGCAAAATGGGAGTCAAGA 1080
    1021 tgtggacagctcccaactcagctcagctcagctcagctcagctcagctcagctcagctc 1080
    1081 TGACAGACCCGATGAAAGTGGCTCGCGCTTGGGAGAGAGCAAAAGCAAACTTAACATGA 1140
    1081 tgacagacccgatgaaagtggctcgcgcttgggagagagcaaaaagcaaaacttaacatga 1140
    1141 ACTATGACAAACTCAGCCGCTGCACTTCCGCTACTACTATGACAAAATATTAAGTAAAG 1200
    1141 actatgacaaactcagccgctgcaacttccgctactactatgacaaaatattagctaaag 1200
    1201 TTGATGTTAAAGCTATGCTTACAAATTTGATTTCCAGGAGATCGCTCAGGCCCTCCAGC 1260
    1201 ttgatgtttaaagctatgcttacaatttgatttccaggagatcgctcaggccctccagc 1260
    1261 CTCACCTCCAGATATCATGATGATCAATATACCAATACCAATACCAATACCAATACCAAT 1320
    1261 ctacacctccagatatacatgatacaataacacacacacacacacacacacacacacac 1320
    1321 ACCATGACACACCCCGAAGATGAATTTGTAGCTCCCATCCCTCCCTGCTTGGCCGTAA 1380
    1321 accatgacacaccccgaaagtgaacttgttagctcccatccctccctgcttggccgttaa 1380
    1381 CCTGATCCAGCTTTTGTGCTGCGCCCTAATCCTACATCTGGAATTCACCAATCGAGCATCT 1440
    1381 cctgatccagcttttgtgctgcccctaatcctacatcagatcagatcagatcagatcagat 1440
    1441 ACCGCAATACGAGGCTGCGAGCTGCTCATATGATGCTCCCATCTTGCGACCTACTATAG 1500
    1441 accgcaatacgaggtgcgagctgctcatatgctcccatcttgcgacctactatag 1500
    1501 TGGGAAAGAAAGAA 1516
    1501 tgggaaagaaagaaa 1516

```

```

RESULT 2
LOCUS      816 bp      mRNA
DEFINITION Erg-3-immunoglobulin heavy-chain enhancer-binding Ets protein
            [alternatively spliced] (mice, pre-B-cell line 22D6, mRNA Partial,
            816 nt).
ACCESSION 566169
NID       566169
KEYWORDS  9436276
SOURCE    Mus sp. pre-B-cell line 22D6.
ORGANISM Mus.
REFERENCE 1 (bases 1 to 816)
AUTHORS  Rivera,R.R., Stuijver,M.H., Steenbergen,R. and Murie,C.
TITLE     Ets proteins: new factors that regulate immunoglobulin heavy-chain
            gene expression
JOURNAL  Mol. Cell. Biol. 13 (11), 7163-7169 (1993)
MEDLINE  94019387
REMARK    Genbank staff at the National Library of Medicine created this
            entry [NCBI glibsg 138523] from the original journal article.
            This sequence comes from Fig. 2.
FEATURES
    1..816
    /organism="Mus sp."

```

gene	/db_xref="taxon:10095"
1. 816	/partial
/note="Immunoglobulin heavy-chain enhancer-binding Ets protein"	
/gene="Erg-3"	
1. 816	/partial
/note="Erg-3"	
/note="Immunoglobulin heavy-chain enhancer-binding Ets protein; This sequence comes from Fig. 2"	
/codon_start=1	
/db_xref="PID:g436277"	
/translation="DDPRLRPSYNADILSHLYHRETRPHLTSDDYKALNSPRLMAHNRNGAARFIRPNTSVYEPATORITRRDLPYEPKRAWGHSLTPQSKAQSPSVPKTEDRPOLDLPOLQIGTPSSRLANGSQIOIMQWLELLSSNSNCITWLGTEDEFKMTDDEVARWMEGRSKSPNNMYDLISRLRYRYNNIMTKVGRARYKFFHGLAQALQPHPPRESSLTKYPSDLPYMYSHAHPOKNFVSPHPALPYTSSSF"	
BASE COUNT	201 a 279 c 172 g 164 t
ORIGIN	
Query Match	36.1%; Score 552; DB 29; Length 816;
Best Local Similarity	84.8%; Pred. No. 0.00e+00;
Matches 691; Conservative 0; Mismatches 121; Indels 3; Gaps 2;	
Db	1 GATGACTCCACGGGCTACGCGGAGCTACATGCCGACATCTCTCTCCACATCTCCAC
Qy	582 gatgactccacgagactcgcgcgagctatacgcagatatctctctgcacacctcac
Db	61 TACCTCAGAGAGACATCCCTTCACATCGATTCGATCCGATGATGAAGCTTTACAA
Qy	642 taccctcagagagactcctcttcacatctgactcagatgagtgtatagaagccttaca
Db	121 AACTCTCCACGGTTAATGATGCTCCGAAACACAGGGGTGACGCTTTATTTCCCAAT
Qy	702 aactctccaggttaatgcatgctgtaaaacacagagagagccacttattttcccaat
Db	181 ACTTCAGTATATCCCGAAGCTACGGAAGATTAACAAGCCGATTAACCTATATGATG
Qy	762 acatcagtttaecccgagaagcaagaaagaaataaacacaaagccagatttaccctatg
Db	241 CCTCCAGAGATCAGCCTGACCGGACACAGCCACACCTCAGCCCTCAGTCCAAAGCTGT
Qy	822 caagagagagatcagctgagcaggtcagcagctacagcca--tc-ccatccagccaagctacc
Db	301 CAGCATATCTCCTCTGCGAGTGCCTCCAAACATGAAGACACAGCTGCTCAGTTAATCCTTAC
Qy	879 caaccatcatcttcaacagctgcccacaaagagaagccagcgtctctcagttgaatcccttat
Db	361 CAGATCTCGGAGCCACACAGTAGCCGCTTGTATATCCAGTAGTAGGGCCAGATCCAGCTG
Qy	939 cagatctcttgagccgacacagcagccgctcttgaaatccaggaagtgaggcagataacagca
Db	421 TGGCAGTCTCTGCTGCAACTCTGTTCAGACAGCTCCAACTCCAACTGCATCAGCTGGGA
Qy	999 tggcagttctctactgagagcttctgctcgagagctccaactccaactgctacacctggag
Db	481 GGCACCAAGGGGAGATTCAAGATGACAGACCGGAGACAGAGTGGCTGCGGCTGGGGGAG
Qy	1059 ggcacacaaatgggagatccaagatgacagaccccttgatgaagtggtctcgagtggtggagag
Db	541 AGGAAGAGCAACCCCAACATGACATGACATGACAGCTCAGCCGCGCCCTCGCTACTACTAC
Qy	1119 aggaagaagcaaaccttaacacttgacatctgacacatccagcgcgtgcaactctgctacactat
Db	601 GACAAAAACATATATACCAAGGTGACAGGGGAAGCGCTACGCTTACAGTTTACTTTCAC
Qy	1179 gacaaaaatatatgactaagaatcataggtlaaacgctatgacctacaaatttgatttccac
Db	661 GGGATGGCCAGAGCCCTGGAGCCCAACCTCTGTAGTGTCTCTGTACAAAGTACCCCTTC
Qy	1239 ggaatcgctcagggcctcagcctcaccctccagaaatcatcatgtacaaataaccatcaa

DB	721	GACCTGCATACATGCGCTCTCTATACACCCACCCACCCAGAGATGAACCTTGTGCTCC	780
OY	1299	gaccctccctcattagcttctaccatgacacacccacagaatgtaacttctgctcc	1358
DB	781	CACCTCCGCGCTCCAGTCACATCTTCCAGTTT	815
OY	1359	catccctctgcttgcctgtaacctcaccagctt	1393
RESULT	3		
LOCUS	HUMERG2	3166 bp ss-RNA	PRI 08-NOV-1994
DEFINITION	Human erg2 gene encoding erg2 protein, complete cds.		
ACCESSION	M17254		
KEYWORDS	erg 2 protein.		
SOURCE	Human cell line COLO 320, cDNA to mRNA, clone lambda 12.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Homiidae; Homo. 1 (bases 1 to 3166)		
AUTHORS	Rao,V.N., Papas,T.S. and Reddy,E.S.		
TITLE	erg, a human ets-related gene on chromosome 21: alternative splicing, polyadenylation, and translation		
JOURNAL	Science 237 (4815), 635-639 (1987)		
MEDLINE	87263429		
FEATURES			
source	Location/Qualifiers		
	1..3166		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/map="21q22.3"		
gene	257..1645		
	/gene="ERG"		
CDS	257..1645		
	/gene="ERG"		
	/note="erg 2 protein"		
	/codon_start=1		
	/db_xref="GDB:G00-119-884"		
	/db_xref="PID:g182187"		
	/translation="MIOIVPDPAAHIKEALSVSEDSOLFCAVGTPLAKTEMTASSSDYGGTSSKSPRVPODDWLSOPARVTKMEGNPSOVNGSRSPDECSVAKGKMVGSPDVGKNGNSYMEKEMPPNMTNRRTVYRPADPLNSTDHROMLENAVKYGLP DVNILFQNDIGKELEKMTKDDFQRLPSTNADILSLHTLRETPLPHLTSDDVDA LQNSPRMLHANNDLPYEPSPRSAMTGHPTPOSKAAPSSTVPTEDORPOLDBY QILPTSTRLANPGSGOIQLOFLELSDSNSSCITWEGTNEFTMTPEDEVARRM GERKSPNNMYDKLSRALRYYYDKNIMKYHGKRAYKDFHGIAOLAPHPPESSLY KERSDLPYMGSYNAHPQKMFVAHPHPALPVYTSSFFFAAPNPYNSPTGCIYPTRLP TSNHPSHLGITY"		
	291..298		
repeat_region	/note="8-bp repeat sequence"		
repeat_region	509..516		
	/note="8-bp repeat sequence"		
polyA_signal	3143..3149		
	/note="putative"		
BASE COUNT	915 a 770 c 768 g 713 t		
ORIGIN	227 bp upstream of PvuII site; Chromosome 21.		
	35.9%: Score 549; DB 26; Length 3166; Best Local Similarity 81.3%: Pred. NMo: 0.00e+00; Matches 1160; Conservative 0; Mismatches 197; Indels 69; Gaps 27;		
DB	289	TATCAAGGAAGCCTTATCAGTTGTGAGTGAGGACCACTGTTGTTTGAATGCTACG	348
OY	74	tattaagaagaacattatcagtggtgaggaagaacagctctgttgggtgctcagg	133
DB	349	AAGCCACACCTGCTTAAGACAGAGATACCGCGTCTCTCCACGACTATGACACAG	408
OY	134	atcgccaccacttgaaagacagaatagacagccctctccagtgatagggcaaac	193
DB	409	TTCCAAGTGAAGCCACGCTCCACACAGATGGCTGTCTCAACCCACGAGGCT	468
OY	194	atcaaatgatgacgcgcgcttcccaacagaagatcggttatacaagccccgcgcaaggt	253
DB	469	CACCATCAAAATGGAATGTAAACCTTAGCCAGGTGATGGCTCAAGAACTCTCTGTGA	528

Oy	254	taccattaaagcttggaagctgtaaccaccaaaccgaagtttaagcttggaagattaccattgatga	313
Db	529	ATGCAGTGTGGCCAAAGCGCGGGAAGATGTGTGGGACGCCCGACACCGTTTGGATGAACTA	588
Oy	314	ctcgacgctggcgaagaagagaggaataatggttaagcagattccagacaattcttggatgaacta	373
Db	589	CGCGACGATCATGAGAGAGAGACATATGCCCACCCCCAAATATGACCCAGAACAGCGCAG	648
Oy	374	tggaaagctaaccttgaaagagaagacatcttcgcgcctccaaataatgaaacaaatgaagaag	433
Db	649	AGTTATCTGTCCACAGACATCTCTACGCTATGGAGTACAGACCATGTGCGCATGTGGGA	708
Oy	434	agttcttctgcacaacagaatccctaaagcttatggagacagaccatgtaacgagctgtgga	493
Db	709	GTGGCGGTGGAAGAATATGTGCGCTTCCAGACGTCAACATCTGTATTCCAGAAACATCGA	768
Oy	494	gtggagcagctgaaagagtaatggtctctccagaagctggaacattctgtcttcagaacaatga	553
Db	769	TGGGAAGCACTGTGCAAGATGACCAAGACGACGACTTCCAGAGGCTCACCCCGACATACAA	828
Oy	554	tgggaagaagatctgttaaaatgaaacaaagatgaactctcagagatccacgscgagctataa	613
Db	829	CGCGACATCTCTTCTCTACACTCTCCACTACCTACAGAAATCTCTTCCACATTTGAC	888
Oy	614	cgcagatctctctctgtcaaccctacacctcctcagaagagatctctcttcacaaattgac	673
Db	889	TTCAATATGATGTGTAAAGCGCTTACAAAAATCTCCAGCGGTAAATGATGATCTTGAAACAC	948
Oy	674	ttcgaatgaatctgtgaataagacctcaaaaatctcccaaggttaatgcaatctctgaacaac	733
Db	949	AG-----A--TTTA-----C-----CAT-A--T-----GA-GCC-C-C-----	968
Oy	734	aggaagagcccaatttaattcttccaaataacatcagattctaacccgaagcaagcaagaat	793
Db	969	--C--CA--G--AGATACGCTG--GA-C---CG---G--TCA-CG--GGC-A--C-C---	100
Oy	794	aacacaaagccagaatttaactcttatggacgaagcagaagatgacgcggtgacgagttacag	853
Db	1002	CCAGGCC-C-C-AGTCCGAAGAGCTGCTCAACACATCTCTTCCACAGTGCCCAAACTGAA	1057
Oy	854	ccattcccaatcagctcaaaagcttaaccacaaacatcatcttcaacagctgcacaaacagaag	913
Db	1060	CCAGGCTCTCAAGTAAATCTTATTCAGATTTCTTGACCAACAAGTACCGCCTTGCAA	1117
Oy	914	ccagacgtctcattgaataatcttcttcagatctcttgcgcgacagacagacgcgttcttcaaa	973
Db	1120	TCCAGGCGAGTGGCAGATATCCAGATTTGGAGTTCCTCTGAGACTCCTGTGGACAGATC	1177
Oy	974	tccagggagctgagcagataaagcctatgacgcttctctctctgtgagctctctgttcgaaacgctc	1033
Db	1180	CAATCTCAGCTGCATTCACCTTGGGAAGCGCCACCAACGGGAGTTCAAGATGACGGATCCGA	1233
Oy	1034	caactccaaatctcatcactccttggagggcacaanaatgaggagtctcaagaatgacagacttga	109
Db	1240	CGAGGTGGCCCGGGCGTGGGGAGCGGGAAGACCAACCAACATGAATACATGAATAGCT	1299
Oy	1094	tgaagctgagctcgagctgcgggagagagaaagaaacctaaacatgaacatgaaacaaat	115
Db	1300	CAGCGCGCCCTTCGTTACTTACTTATGCAAGAACATATCAATACCAAGATCCATGGAGAAGC	1357
Oy	1154	cagcgcgtgcaactctgctactactatgctcaaaaataatataatgaactaaagtctcatgtaag	121
Db	1360	CTAGGCTCAACAAGTTTCGACTTCCACGGGATATGCCAGGCGCTTCCAGCCCAACCCCGGA	1411
Oy	1214	ctatgctctacaataattgattctccacggaaatctgctcaagccctccacgctcactacactcaga	127
Db	1420	GTCATCTCTGACAAAGTACCCCTCAGACATCCCGTATACGGGTCCGTATACGCCCAACCC	1477
Oy	1274	atcatctcatgtacaataatcccatatagaacctctcccttaaatgattctccacaatgacaccc	133
Db	1480	ACAGAAGATGAATTTGTGGCGGCCCAACCTCCAGGCCCTCCCGTGACATCTTCCAGTTT	1533

Query	1334	cgagaagatgaacttctgtagctcccccacccctctgttgcgcgtaactcatccaagctt	1393
Db	1540	TTTTGCTGCCCCAAACCATAGTGAATTCACCAATGAGGGGTATATACCCACACATAG	1599
Qy	1394	ttttctgcccataactcatctagtaattcaccaactgagagcatctaccacaatcaag	1453
Db	1600	GCTCCCAACAGCCATATGCTTCATCTGGGCACTACTACTAA	1645
Qy	1454	gtctgcagctctgctcatatgctctcccatcttgagcactactactaa	1499

RESULT#	4	HUMERGI1	3126 bp	mRNA	PRI	15-MAR-1989
LOCUS						
DEFINITION		Human erg protein (ets-related gene)	mRNA, complete cds.			
ACCESSION		M21535	M17390			
NID		g182182				
KEYWORDS		erg protein.				
SEGMENT		1 of 2				
SOURCE		Human, cell line COLO 320, cDNA to mRNA, lambda-7.				
ORGANISM		Homo sapiens				
REFERENCE		Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Euthera; Primates; Catarrhini; Homiidae; Homo.				
AUTHORS		Reddy,E.S.P., Rao,Y.N. and Papas,T.S.				
TITLE		The erg gene: A human gene related to the ets oncogene				
JOURNAL		Proc. Natl. Acad. Sci. U.S.A. 84, 6131-6135 (1987)				
MEDLINE		87317608				
FEATURES						
source		location/Qualifiers				
		1..3126				
		/organism="Homo sapiens"				
		/db_xref="taxon:9606"				
		195..1286				
		/note="erg1 protein"				
		/codon_start=1				
		/db_xref="PID:g182185"				
		/translation="MVGSPDTVGNNTSGSYEEKMPPPNMTTNERRYIVPADPTLNTST				
		DHVRQWLEMAVKEYGLPDVNTILFQNTDGELEKMTDDQRLPSYNADILSLHY				
		LRETPHLTSDSDPKALQNSPLMHAINTDLPYEPERSAMTGHGPTQSKAOPS				
		LSYPTKEDORPOLDPOILGPTSSRLANPESGQIOLOMOFLLELLSSNSSCITWGG				
		TNGEFKMTDDEVAARMGERKSRKMYNMDLSRLRYVYKNTKYHGKRYAKKPP				
		HGIALQAPHPRESLKYRPSDLPEKMSYHAPKMFVAPHPALPYTSSFFAAPNP				
		PYNSPTGGIYPTNRLPTSHMPSHLGly"				
BASE COUNT		928 a 732 c 725 g 741 t				
ORIGIN		1 bp upstream from EcoRI site.				

Query Match	30.3%	Score 463;	DB 26;	Length 3126;
Best Local Similarity	84.2%	Pred. No. 0.00e+00;		
Matches 588;	Conservative 0;	Mismatches 107;	Indels 3;	Gaps 1;
Db	589	CAGATTACCATATGAGCCCCCAGAGAGATGACGCTGACGGGTACAGCCACCCACGC	648	
Qy	805	cagattaccattatgagcaagcgagagatcagctgtgagcagatcaacgcatcccatc	864	
Db	649	CCCACTCGAAGCTGCTCAACCATCTCCTTCACAGTCCCAAAATGTAAGACCAAGCTC	708	
Qy	865	--agtc aaagctacccacacatcatcttcaacagtgcccaaaacagagaccagctg	921	
Db	709	CTCAATTGATCTTATAGATTCTTGACACCAAGAATAGCGGCTTGCAATTCAGGCA	768	
Qy	922	ctcagttgatcttcatatgattcttgcagcagcagcagcagcgtcttgcaatcaga	981	
Db	769	GNGCCAGATCCAGGTTTGGCAGTTCCTCCCTGAGACCTCCTGCGGACGATCCAACTGCA	828	
Qy	982	gtgggcagataacagtaicagctacgtctccacagagatctctgtggaacgctccaatcca	1041	
Db	829	GCTGCATCACTGGGAAGGACCAACAGGGGAGTTCAAGATACGGATCCGACGAGTGG	888	
Qy	1042	actgatacactctggagggcacaatgtgggttcagaatgaaccctgtatgaagcgtg	1101	
Db	889	CCCGGCGCTGGGGAGCGCGGAAGCAAAACCAACATGACTAGATTAAGCTCAGCGCG	948	
Qy	1102	ctcgcgcgtctgggagagagaaagcaaaactaacatgaactatgaacaaactcugcgtg	1161	

RESULT	5	762 bp	mRNA	PRI	24-JAN-1995
LOCUS	S72621				
DEFINITION	EMS. . . erg (translocation, type 1e and 9e) [human, SK-PN-LI cell]				
ACCESSION	S72621				
NID	9633773				
KEYWORDS					
SOURCE	human SK-PN-LI cell line.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryote; Mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Carnivhni; Homnidae; Homo. 1 (bases 1 to 762)				
AUTHORS	Giovannini,M., Biegel,J.A., Serra,M., Wang,J.Y., Wei,Y.H., Mycum,L., Emanuel,B.S. and Evans,G.A.				
JOURNAL	EMS-erg and EMS-Fill fusion transcripts in Ewing's sarcoma and primitive neuroectodermal tumors with variant translocations				
MEDLINE	J. Clin. Invest. 94 (2), 489-496 (1994)				
REMARK	GenBank staff at the National Library of Medicine created this entry [NCBI gidsb 153609] from the original journal article. This sequence comes from Fig. 2A.				
FEATURES	Map location: t(21;22)(q22;q12).				
source	Location/Qualifiers				
gene	1..762				
gene	/organism="Homo sapiens"				
gene	/db_xref="taxon:9606"				
gene	1..762				
gene	/partial				
gene	/note="fusion gene"				
gene	1..85				
gene	/partial				
gene	/gene="EMS"				
gene	1..762				
gene	/partial				
gene	/gene="EMS-erg"				
gene	/note="This sequence comes from Fig. 2A."				
gene	/codon_start=1				
gene	/product="EMS-erg fusion protein type 1e"				
gene	/db_xref="PID:g633774"				
gene	/translation="PTSPYPTGTSYQADSPQSYQSSSYSGQNLPIEPPRRASMTGHPSPQKAAQSPSTVPKTDQRPOLDPYQILGPTSSRLANPDSQIQLOMFLLELHSPSSICITWEGNGEFKMTDPDEVARRMGERKSKPMNYDKLSRLALRYYKRNLEL VHGRSYRKFDFHIAQLOLPHPESSSLKYPEDLPYMSYHNPQKNFVAHPHPGVSSFEAPANPNYMSPTGCIYPTNRLPTSHMPSH"				
gene	86..762				
gene	/partial				
gene	/gene="erg"				

BASE COUNT	191 a	272 c	162 g	137 t
ORIGIN				
Query Match	29.2%; Score 446; DB 26; Length 762;			
Best Local Similarity	84.0%; Pred. No. 0.00e+00;			
Matches	569; Conservative 0; Mismatches 105; Indels 3; Gaps 1;			
Db	86	ATTATACATATGAGCCCCCAAGAGATCAGCTTGAGCCGGTCACGCGCACCCACGCC	145	
Oy	808	atttcctctatgagcaagcagagatgaatcagctgagagctcagacagcagccacac	864	
Db	146	AGTGAAGCGCTCAACCATCTCCTTCCACAGTGGCCCAAACTGAGACACGCTCCTC	205	
Oy	865	agtcaaaagcaccacaacacatcattcaaaagtgcccaaaagaaagaccugcttc	924	
Db	206	AGTTGATCTTTATCAGATTTCTTGAGACCAACAATATACCGCTTGTCAAAATCAGCAGT	265	
Oy	925	agttgatcctctatcagattctctgagccgagcagcagcgcgtctgcaaatccagggagt	984	
Db	266	GCCAGATCCAGCTTTGGCAGTTCCTCGTAGCTCTGTCTGGACAGCTCCAACTCAGCT	325	
Oy	985	gscagatcacagctatgtagcttctctactgagcttctgctgagacagctccaatcc	1044	
Db	326	GCATCACTGGGAAGGCCCAACGGGAGTTCAAGATGACGATCCCGACAGAGTGGCCC	385	
Oy	1045	gcataccttggagggcgaataatgggagttcaagaatgacagaccctgataagtgctc	1104	
Db	386	GGCGCTGGGAGAGCGGGAAGAGCAAAACCAACATGAACTACGATAGCTCAGCGCGCC	445	
Oy	1105	ggcgtctgggagagagagaaagcaaacctaacatgaaactaagaaactcagcgtgac	1164	
Db	446	TCCGTTACTATATGACAAGAATCATGACCAAGTCCATGGGAGCGCTACGCTTACA	505	
Oy	1165	ttcgctactactatgacaaaatatattagtaaatgattcatgtaaaagctatgctaca	1224	
Db	506	AGTTCGACTTCCAGGGATCGCCCAAGGCGCTTCAGGCCACCCCGGAGTCACTCTGT	565	
Oy	1225	aattgattcttcacggaaatgctcagagccctcagcctcacctccagaatcatcatt	1284	
Db	566	ACAAGTACCCCTTCAGACCTCCCGTACATGGGCTCTTATCAGCGCCACCAAGAAATGA	625	
Oy	1285	acaaataccatcacgaacctccctccatcatgagttcttcacatgacaccccagaaatga	1344	
Db	626	ACTTGTGGCGCCACCCCTTCAGACCCCTCCCGGTAGACATCTTCAGATTCTTGTGCTGCC	685	
Oy	1345	acttgtagctcccatccctccctgctttgcccgaacctcatccagctttttgtgccc	1404	
Db	686	CAAACCCATCTGGAATTCACCAACTGGGGGTATATACCCACACTAGAGCTCCCAACA	745	
Oy	1405	ctaaccatactggaattcaccaactgagagcatcaccacaatacagagctgscagctg	1464	
Db	746	GCCATATGCTTTCAT 762		
Oy	1465	ctcatagctcttcccat 1481		
RESULT	6			
LOCUS	AP000022	133746 bp	DNA	PRI 13-MAY-1998
DEFINITION	Homo sapiens genomic DNA, chromosome 21q22.2 (Down Syndrome			
	region), segment 15/15, complete sequence.			
ACCESSION	AP000022			
NID	g3132332			
KEYWORDS	HTG.			
SOURCE	Homo sapiens DNA.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;			
	Primates; Catarrhini; Hominiidae; Homo.			
REFERENCE	1 (bases 1 to 133746)			
AUTHORS	Hirakawa, M., Yamaguchi, H. and Imai, K.			
TITLE	Homo sapience 1,296,826bp genomic DNA of 21q22.2 Down Syndrome			
JOURNAL	region			
REFERENCE	Published Only in Database (1998) In press			
	2 (bases 1 to 133746)			

LOCUS	8	CCRNALI	3490 bp	mRNA	VRT	07-JUL-1998
DEFINITION		Coturnix coturnix mRNA for transcription factor FLI, clone 10.1.				
ACCESSION		Y14773				
NID		93269302				
KEYWORDS		fli gene; FLI oncoprotein; transcription factor.				
SOURCE		quali.				
ORGANISM		Coturnix coturnix				
REFERENCE		Eukaryota; Metazoa; Chordata; Vertebrata; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Coturnix.				
AUTHORS		1 (bases 1 to 3490) Mager,A.M., Grapin-Botton,A., Ladfall,K., Meyer,D., Wolff,C.M., Stiegler,P., Bonnin,M.A. and Remy,P.				
TITLE		The avian fli gene is specifically expressed during embryogenesis in a subset of neural crest cells giving rise to mesenchyme				
JOURNAL		Int. J. Dev. Biol. 42 (4), 561-572 (1998)				
MEDLINE		98358003				
REFERENCE		2 (bases 1 to 3490) Stiegler,P.				
AUTHORS		Direct submission				
TITLE		Submitted (07-SEP-1997) P. Stiegler, Centre National de la Recherche Scientifique, U.P.R. 9005 du C.N.R.S., 15 Rue Rene Descartes, 67084 Strasbourg Cedex, FRANCE				
JOURNAL		Related sequence Y14774.				
COMMENT		Location/Qualifiers				
FEATURES		1..3490				
SOURCE		/organism="Coturnix coturnix"				
		/db_xref="taxon:9091"				
		/chromosome="1"				
		/dev_stage="4-day embryo"				
		/clone_1lb="pcDNA1"				
		/clone="10.1"				
		/map="q31-q3"				
		215..1513				
		/gene="fli1"				
		215..1513				
		/gene="fli1"				
		/gene="fli1"				
		/function="transcription factor"				
		/codon_start=1				
		/product="FLI transcription factor"				
		/db_xref="PID:e1308265"				
		/db_xref="PID:g3269303"				
		/translation="MDGTIKELALSVSDDSLPSTYGAATHLPKADMTASGNDPYGCA PKHINPLPPOEWINOPVRYVREKYDGHMMSRSPVDCSVNKSILVGAGTESPNPMS XYTMDEKNGPRPMNTNERVYIPADPLMTQOHYOMLEMAKEYGEMIDTTFPG MMDKEICKMKKIDPLRTSLYINTEVLISHSLYRESSLAIATPSTIENSRLATK EGPVAAGTQNVNKTTEQRPQPDYQILGPTSSLANGSGQIDLMQFLLELSDSSN ASCITWGEATNNEFMTDPPDEVARRMGRKSKPNNNYDKLSPALRYDYDKNIMTVHGRK RYAKCFDGHGIALQLOPHPTESSMYKYPSPDLSYKPSYHAHQKVNFVPPHPSMPTVTS SSFGAASPWTSFSPAGSIYPPNPVPRHNMNVSHLSLSY"				
BASE COUNT		1026 a 770 c 752 g 942 t				
ORIGIN						
Query Match		20.9%; Score 320; DB 21; Length 3490;				
Best Local Similarity		79.2%; Pred. No. 4,49e-263;				
Matches		434; Conservative 0; Mismatches 114; Indels 0; Gaps 0;				
Db	910	AACAAACAGCAGACGGCCCCAACACAGATTCATATCAATCTCTGGGGCCACCAAGTAGTCG 969				
Qy	905	aacaagaacacagcgctccctcagttagatcctatcagattcttggacagcagcagcgcg 964				
Db	970	TCTTCGCATCCTGTGGAGTGGGAGATACAACTGTGGGAGTCCCTCTCGAGATTGCTGTC 1029				
Qy	965	tcttgcacaatccagggagattgggcagataacagctatggcaggtccctcctggagctctg 1024				
Db	1030	GGACAGTTCCAATGCCAGCTGTATGATACATGGGAAGGACCAATGGGGAATTCAGATGAC 1089				
Qy	1025	ggacagctcccaactcaactgcatcactgcctggaggagcacaatgggaggttcaagatgac 1084				
Db	1090	AGACCCAAATGAATGGCGACGGCGCTGGGGGAGAAACGCAAAAGCAGCCCAATATAATTA 1149				
Qy	1085	agaccctatgaagtgagctcgcgcttggaggagagaggaagcaaacctcaactcaactgaac 1144				

LOCUS	9	CCFLIONCO	3545 bp	mRNA	VRT	07-JUL-1998
DEFINITION		Coturnix coturnix	mRNA for transcription factor FLI, clone 6.1.			
ACCESSION		Y14774				
NID		g3269304				
KEYWORDS		fl1 gene; FLI oncoprotein; transcription factor.				
SOURCE		quail				
ORGANISM		Coturnix coturnix				
		Eukaryota; Metazoa; Chordata; Vertebrata; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Coturnix.				
REFERENCE		1 (bases 1 to 3545)				
AUTHORS		Mager, A.M., Graplin-Botton, A., Ladjail, K., Meyer, D., Wolff, C.M., Stiegler, P., Bonnin, M.A. and Remy, P.				
TITLE		The avian fl1 gene is specifically expressed during embryogenesis in a subset of neural crest cells giving rise to mesenchyme				
JOURNAL		Int. J. Dev. Biol. 42 (4), 561-572 (1998)				
MEDLINE		98358003				
REFERENCE		2 (bases 1 to 3545)				
AUTHORS		Stiegler, P.				
JOURNAL		Direct Submission				
TITLE		Submitted (09-SEP-1997) P. Stiegler, Centre National de la Recherche Scientifique, U.P.R. 9005 du C.N.R.S. 15 Rue Rene Descartes, 67084 Strasbourg Cedex, FRANCE				
COMMENT		Related sequence Y14773.				
FEATURES		location/qualifiers				
source		1..3545				
		/organism="Coturnix coturnix"				
		/db_xref="taxon:9091"				
		/chromosome="1"				
		/dev_stage="4-day quail embryo"				
		/clone.lib="pcdN1"				
		/clone="6.1"				
		/map="q31-33"				
		396..1595				
		/gene="fl1"				
		396..1595				
		/db_xref="PID:el308263"				
		/db_xref="PID:g3269305"				
		/translation="MTASGNPDYGGPCHINDLPDQEMINOPVAVNYKREYDHNGSR				
		ESPDCVNRCSKLVAGCTESNPMKSYNDDEKGPNNMTNRRIVADPTIMQ				
		EHVQKLEWAKETGLMEIDITFPQNDGKELCKMKNKDDLPRTSLINTEVYLSHST				
		LRSSSSLLAVTTPSHTEASSKLATKEGPPVAGTGVNKKTTTQQPQPPQVILGPISSN				
		RLAPGSGQIDLMQFLLELSDSSNACIWEGTNGEFGKTIIDPEVARRNGERSKRN				

	misc_feature	MNNDKLSRALRYVYDKDNIIMRVHGKRARAYFEDDFGIALOAPHETESMKKKYPSPDSL MPSTHAHQOKYNVFPPHPSSMPTVTSSSFCAASPYWTSPAGSITPNNPNVRHNPAAVS
	PHLGSYY"	
	1071..1320	
	/gene".fl1"	
BASE COUNT	1002 a 803 c 768 g 972 t	
ORIGIN	"note"Ets binding domain"	

Query Match	Similarity	Score	DB	Length
Best Local	79.0%	Pred. No. 3,53e-261;		
Matches	433; Conservative	0; Mismatches 115; Indels 0; Gaps 0;		
Db	992 AACAGACAGACAGAGGCCCCAACACAGATCCATATCAATCTCTGGGGCCACCACTAGTCG	1051		
Qy	905 aacgaagaacacagctcctcagctcagctcagctcagctcagctcagctcagctcagctcagctc	964		
Db	1052 TCTTGGCCAATCCTGGGAGTGGGACAGATACAACTGTGGCAGTTCTCTCGAGTTGCTGT	1111		
Qy	965 tcttcgaatccagggagctggcagatcagctcagctcagctcagctcagctcagctcagctcagctc	1024		
Db	1112 GGACAGTTCCTCAATGCGCAGCTGTATCATCATGTGGGAAGGACACATGGGAATTCAGATGAC	1171		
Qy	1025 ggaacagctcccaactccaaactcgtcaccctggagggcacaatctggggagttcagaatgac	1084		
Db	1172 AGACCACATGAATAGGCGAGCGCTGGGGGGAACGGCAAGGCCAAGCCAAATGATTTA	1231		
Qy	1085 agaccctgaatgaatgagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct	1144		
Db	1232 TGACAGCTGAGCGGAGCGCTTCATCATCTACTAGATTAAGAACTTATGACCAAGTCA	1291		
Qy	1145 tgcacaactcagccgctgactcgtcactactactactgacaacaatatctgactaaagtcca	1204		
Db	1292 TGCCAAAGATGATGCTCAATTAATTTGACTTTCATGCGATGGCCAGGCCCTCCAGCTCA	1351		
Qy	1205 tgttaaaagctatgctcctaacaatttgatttcacaggaatcgctcagccctccagctcca	1264		
Db	1352 TGCCACTGAATCGTCGATGTACAGTATCCGTGAGATCTCTCTACATGCTTCTTACCA	1411		
Qy	1265 ccttcacgaatcatcatcatgataaatacccaaccacagactccctcactatgagttcactca	1324		
Db	1412 TGCCACACAGAGAGAAGGTGAATTTTGAACCCCAACACCTCTCTTATGCTGTACATC	1471		
Qy	1325 tgcacaacccccagagaatgaaacttctgtgcccaccacacccctcctgtcccgaaacctc	1384		
Db	1472 ATCCAGTTCCTTTGGAGCAGCGCTCACCCTTATTTGGACCTCCCTCTGGAACATTTATCC	1531		
Qy	1385 atccagctcttttctgtcgtccctataccatctgtaattcacacaactcgtgaggcactacc	1444		
Db	1532 AAACCCCA 1539			
Qy	1445 caatacca 1452			

[illegible]

QY	1392	ttttttgtgcacctaatccatacttggaattccaccacactgagggcactctaccataacc	1451
Db	1547	AG 1548	
QY	1452	ag 1453	
RESULT	11		
LOCUS	A36461	2916 bp	DNA
DEFINITION	Sequence 2 from Patent WO9323549.		PAT
ACCESSION	A36461		
NID	g2293779		
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Carnivora; Homiidae; Homo.		
AUTHORS	1 (bases 1 to 2916)		
TITLE	Aurias, A., Delattre, O., Desmaziere, C., Melot, T., Peter, M., Plougastel, B., Thomas, G. and Zucman, J.		
JOURNAL	NUCLEIC ACID CORRESPONDING TO A GENE OF CHROMOSOME 22 INVOLVED IN RECURRENT CHROMOSOMAL TRANSLOCATIONS ASSOCIATED WITH THE DEVELOPMENT OF CANCEROUS TUMORS		
COMMENT	Patent: WO 9323549-A 2 25-NOV-1993; CENTRE NAT RECH SCIEN (FR) Other publication FR 2691475 931126 Other publication JP 85009647 960206.		
FEATURES	Location/Qualifiers		
BASE COUNT	1..2916		
ORIGIN	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/dev_stage="ADULT"		
	844 a 686 c 663 g 723 t		
Query Match	20.4%; Score 312; DB 22; Length 2916;		
Best Local Similarity	78.8%; Pred. No. 1,71e-255;		
Matches	427; Conservative 0; Mismatches 115; Indels 0; Gaps 0;		
Db	882	AGAGCAACGGCCCCAGCCAGATCCGATTCAGATCTTGCGCCCGACGACGATCGCTTACG	941
QY	911	agaccagcgctctcagtagaactctctacagattcttgcgacgacgacgacgacgcttgc	970
Db	942	CAACCTGAGGAGCGGGGAGATCAAGCTGTGGCAATTCCTCTCGAGAGTGTCTCCACAG	1001
QY	971	aaattcaggagatgagagatgacacatgagcttctctctgagcttctgctgagacag	1030
Db	1002	CGCCAAACGCCAGCTGATACACTTGGAGGGGACCAACGGGAGATTCAAAATGACGAGCC	1061
QY	1031	ctccaactccaactgcatcaccttgctgagggagacaaatgaggagattcagaatgacagacc	1090
Db	1062	CGATGAGGTGGCCAGGCGCTGGGGCGAGCGGAAAAAGCAAGCCCAACATGAAATACGACAA	1121
QY	1091	tgatgaagtgcgctgcgctgctgagggagagagaaagaacacactacatgaaactatgacaa	1150
Db	1122	GCTAGCGCGGCGCCCTCGCTTATCTACTGATGATAAAACATATATATAGCAAAATGACAGGCA	1181
QY	1151	aactcagcgctgacttgcctactctacttgaacaaaataataataagtaactaaagtgttaa	1210
Db	1182	AAGATATGCTTCAAAATTTGACTTCCACAGCGATGCGCCAGGCTCTGACGACATCGAC	1241
QY	1211	agcctatgctcacaataattgatttccacaggaatgcttccagggccctccagctcactcc	1270
Db	1242	CGAGTGTTCATGTACAAAGTACCCTTGACATCTCTTACATGCGCTTCTTACATCGCCA	1301
QY	1271	agaatcattcattatgacaaataccatcagacctccctacatgattctcattacatgacaa	1330
Db	1302	CCAGCAGAAAGGTGAATTTGCTTCCCTCCCAACCAATCTCCATGCTGCTGCTGCTGCTGCTG	1361
QY	1331	cccccaagaagatgaaatttctgacttcccccaatccccctgcttgcgcgttaacctcaatccag	1390

Db	1362	CTTCTTGAGACCCATCACATATCTGGACCTCCCGGGGGAATCTACCCCAACC	1421
Qy	1391	cttttgcgcctccatcatcactcgaattcaccaactcgaagcatctaccccaatac	1450
Db	1422	CA 1423	
Qy	1451	ca 1452	
RESULT	12		
LOCUS	HSNUMFLI	2938 bp	RNA
DEFINITION	H.sapiens HUMFLI-1 mRNA.		PRI
ACCESSION	X67001	544250	28-JUN-1995
NID	932529		
KEYWORDS	FLI-1 gene homologue.		
ORGANISM	human.		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
JOURNAL	1. (bases 1 to 2938)		
TITLE	Delattre, O.		
JOURNAL	Direct Submission		
TITLE	Submitted (26-MAY-1992) O. Delattre, Lab. de Genet. des Tumeurs.		
REFERENCE	Inst. Curie, 26 rue D'Ulm, 75231 Paris Cedex, FRANCE		
AUTHORS	2. (bases 1 to 2938)		
	Delattre, O., Zucman, J., Plougastel, B., Desmaza, C., Melot, T., Peter, M., Kovar, H., Joubert, I., de Jong, P., Rouleau, G., Aurias, A. and Thomas, G.		
	Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours		
	Nature 359 (6391), 162-165 (1992)		
FEATURES	92396239		
SOURCE	Location/Qualifiers		
JOURNAL	1. .2938		
MEDLINE	/organism="Homo sapiens"		
FEATURES	/db_xref="taxon:9606"		
SOURCE	/dev_stage="adult"		
	/tissue_type="bone marrow"		
	/clone_id="cDNA, Clontech HLI058"		
	/chromosome="11q24"		
	143. .1501		
gene	/gene="HUMFLI-1"		
	143. .1501		
CDS	/gene="HUMFLI-1"		
	/codon_start=1		
	/product="homologue of the murine FLI-1 gene"		
	/db_xref="PID:g32530"		
	/db_xref="SWISS-PROT:Q01543"		
	/translation="MDGIRKALSVSDDSLFDSAYGAHLPRKADMTASGSPDYGO		
	PHINPLPQOEWINOPRYNVKREVYDHNGMSRSPVDCSVSKSLVGGESNPMNY		
	NSYMDENCKRPPTNNTERRIVPADPTLMTDEHYQWLEMAIKESLSEIDISFQ		
	MDGKEKCKRNKEDFLRATLITNTLVLSLSTLRESLLAINTSHSDSSRLSVRE		
	DPSEYVRRGAGWNMNSGLTKSPPLGAGQITKNTGROQPDYQILGFTSSRLAP		
	GSQIQLMOFLLELSDSANASCIITWGTGTEFKMDPDPEVARMGSRKPNNYDK		
	LSLRATYDKNNTKMGKRYAKYKFFHIAALQGPRESSMYKPSDISYSPYVH		
	AHOKVNFVPHRSPMSMPTSSFFGAASQWTSPGTGYNPVPRNPVHVPVHLS		
	YV"		
misc_feature	971. .1222		
	/gene="HUMFLI-1"		
	/note="ETS domain"		
polya_signal	2390. .2395		
polya_signal	2411. .2416		
polya_signal	2415. .2420		
polya_signal	2908. .2913		
	/evidence=experimental		
polya_site	2927		
BASE COUNT	847 a 692 c 675 g 724 t		
ORIGIN			
Query Match	20.4%.	Score 312;	DB 26; Length 2938;
Best Local Similarity	78.8%;	Pred. No. 1,74e-255;	
Matches 427; Conservative	0;	Mismatches 115;	Indels 0; Gaps 0;

Db 904 AGAGCAAGCGCCCGCCAGATCCGATATCAGATCTCTGGGCGCCGACGACGAGTCCGCTAGC 963
Qy 911 agaccagcgctccctcagcttagaaccctacgaattctctgagccagcagcgctctgc 970
Db 964 CAACCCCTGAAGCGGCGAGATCCAGCTGTGCAATTCCTCGAGCTCCTCCGACAG 1023
Qy 971 aaatccaggaggtgagcagatcagctagtcagctcctcagcgtctctctcgagacg 1030
Db 1024 CGCCAAAGCCAGCTGTATCAGCTGGAGGGGACCAAGCGGAGTTCAAAATGACGGACCC 1083
Qy 1031 ctccaaactccaactgcatacacttgagagagcaaatgggaggttcaagatgacagacc 1090
Db 1084 CGATGAGTGGCGACGCGCTGGGCGGAGGAAAGCAAGCCCAATGATTAAGACAA 1143
Qy 1091 tgatgaagctgctgcgctgctgggagagagaaagcaaaccttaacgaactaagacaa 1150
Db 1144 GCTGAGCGCGGCGCTCCGTTATCTATGATATAAATGATTAAGCAAGTGCAGGCAA 1203
Qy 1151 actcagcgctgactctgactacacaaagcaaaatataagactaaagttcaatgtaaa 1210
Db 1204 AAGATATGCTTACAATTTGACTTCCACGGCATTTGCGCCGCTTGACGACCATCCGAC 1263
Qy 1211 agcgtatgctcctcaaatctgattctcaggaatcgctcagcgccctccagcctccac 1270
Db 1264 CGAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1323
Qy 1271 agaatcattcctgtacaaatcctacacacacacacacacacacacacacacacacac 1330
Db 1324 CGAGCAGAAGGTGAAGTGTGCTTCCCTCCCATGCTCCATGCTCCATGCTCCGAC 1383
Qy 1331 ccccaagagaatgaaactgtgtgctccacacacacacacacacacacacacacacacac 1390
Db 1394 CTTCTTTGAGCGCGCATCAATTAAGTGAACCTCCCGACGGGGGAGATCTACCCAAACC 1443
Qy 1391 ctttttgcgcctcctacacacacacacacacacacacacacacacacacacacacacac 1450
Db 1444 CA 1445
Qy 1451 ca 1452

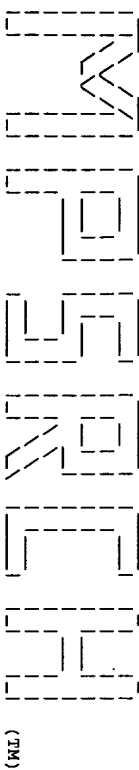
RESULT 13
LOCUS 545205 1673 bp mRNA PRI 05-JAN-1993
DEFINITION FLI-1-Friend leukemia integration 1 [human, mRNA, 1673 nt].
ACCESSION 545205
NID 9257353
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Prasad,D.D., Rao,V.N. and Reddy,E.S.
TITLE Structure and expression of human FLI-1 gene
JOURNAL Cancer Res. 52 (20), 5833-5837 (1992)
MEDLINE 93007976
REMARK GenBank staff at the National Library of Medicine created this entry [NCBI gidsq 115336] from the original journal article.
FEATURES
SOURCE location/Qualifiers
1..1673
gene /organism="Homo sapiens"
/db_xref="taxon:9606"
134..1492
/note="Friend leukemia integration 1"
CDS /gene="FLI-1"
134..1492
/gene="FLI-1"
/note="Friend leukemia integration 1; This sequence comes from Fig. 1A"
/codon_start=1
/db_xref="PID:9257354"
/translation="MDGTRKALSVSDDSLFPSAYGAHAHLPRKMTASGSPDYGO

BASE COUNT 428 a 488 c 429 g 328 t
ORIGIN
Query Match 20.3%; Score 310; DB 26; Length 1673;
Best Local Similarity 78.6%; Pred. No. 1,346-233;
Matches 426; Conservative 0; Mismatches 116; Indels 0; Gaps 0;
Db 895 AGAGCAAGCGCCCGCCAGATCCGATATCAGATCTCTGGGCGCCGACGACGAGTCCGCTAGC 954
Qy 911 agaccagcgctccctcagcttagaaccctacgaattctctgagccagcagcgctctgc 970
Db 964 CAACCCCTGAAGCGGCGAGATCCAGCTGTGCAATTCCTCGAGCTCCTCCGACAG 1014
Qy 971 aaatccaggaggtgagcagatcagctagtcagctcctcagcgtctctctcgagacg 1030
Db 1015 CGCCAAAGCCAGCTGTATCAGCTGGAGGGGACCAAGCGGAGTTCAAAATGACGGACCC 1074
Qy 1031 ctccaaactccaactgcatacacttgagagcaaatgggaggttcgaatgacagacaccc 1090
Db 1075 CGATGAGTGGCGACGCGCTGGGCGGAGGAAAGCAAGCCCAATGATTAAGACAA 1134
Qy 1091 tgatgaagctgctgcgctgctgggagagagaaagcaaaccttaacgaactatgaa 1150
Db 1135 GCTGAGCGCGGCGCTCCGTTATCTATGATATAAATGATTAAGCAAGTGCAGGCAA 1194
Qy 1151 actcagcgctgactctgactacacacacacacacacacacacacacacacacacacacacac 1210
Db 1195 AAGATATGCTTACAATTTGACTTCCACGGCATTTGCGCCGCTTGACGACCATCCGAC 1264
Qy 1211 agcgtatgctcctcaaatctgattctcaggaatcgctcagcgccctccacacaccc 1270
Db 1255 CGAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1314
Qy 1271 agaatcattcctgaatcctacacacacacacacacacacacacacacacacacacacacac 1330
Db 1315 CGAGCAGAAGGTGAAGTGTGCTTCCCTCCCATGCTCCATGCTCCATGCTCCGAC 1374
Qy 1331 ccccaagagaatgaaactgtgtgctccacacacacacacacacacacacacacacacac 1390
Db 1375 CTTCTTTGAGCGCGCATCAATTAAGTGAACCTCCCGACGGGGGAGATCTACCCAAACC 1434
Qy 1391 ctttttgcgcctcctacacacacacacacacacacacacacacacacacacacacacac 1450
Db 1435 CA 1436
Qy 1451 ca 1452

RESULT 14
LOCUS HUMFLI1A 1932 bp mRNA PRI 31-DEC-1994
DEFINITION Human FLI-1 mRNA, complete cds for two alternate splicings.
ACCESSION M93255
NID 9182659
KEYWORDS
SOURCE Homo sapiens (tissue library: lambda g11 HEL 2T) blood cDNA to
factor.
ORGANISM Homo sapiens
Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Thomas,R., May,W., Denny,C., Raskind,W., Moore,J., Maki,R.A.,
Beck,E. and Klemenz,M.J.
TITLE Human FLI-1 localizes to chromosome 11024 and has an aberrant
transcript in neuroepithelioma

Db	1114	CGATGAGSTGGCCAGGCGCTGGGGGCGAGCGGAAACACGACCCCAACATGAAATTACGACAA	1173
Qy	1091	tgaTgaagTggctctcggtcttggggagagagaaagcaacctaaactgaactgaactgaTgaaca	1150
Db	1174	GCTAGCGCGGGCCCTCCGCTATTACTGATGATATAAAACATTATGACCAAAAGTGCACGGCAA	1233
Qy	1151	acttaagccgtgcactctgcctactactatgacaaataataTgaactaaagtTcatgTga	1210
Db	1234	AAGATATGCTTACAAATTTGACTTCACGCGCAATGGCCAGGCTTCGACGCCACATCCGAC	1293
Qy	1211	acgtatgacctaaaatTtgaattccaagaaTcgtctgaagccctccagcctccacccctc	1270
Db	1294	CGAGTCTGCATGTACAAAGTACCCCTTCTGACATCTCTACATGCTTCCCAACATGGCCA	1353
Qy	1271	agaatcatcatctgTgaacaaataccaTcaagcctccctcatagTgTctctatacTgaaca	1330
Db	1354	CCACAGAAAGSTGAATCTTTGTCCCTCCGCAATCATCTTCATGCGCTGTCACTTCTCCAG	1413
Qy	1331	cccccaagaagatTgaactTtgaactcccccactccctcgtcttgcgcgTaaacctcaTccag	1390
Db	1414	CTTCTTGGAGCGCATCACATATATGGA	1442
Qy	1391	ctttctctgTgcctctaaTcccaTcaTccTgga	1419

Search completed: Sat Apr 10 12:38:54 1999
Job time : 2833 secs.



Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MSearch - protein database search, using Smith-Waterman algorithm

Run on: Wed Apr 7 09:43:13 1999; Maspar time 20.29 Seconds
882,246 Million cell updates/sec

Output not generated.

Title: >US-08-878-177-4
Description: (1-478) from US08878177.pep
Perfect Score: 3467
Sequence: 1 MASTIKREALSVSEDSLFE.....IYPNRLPAHMPSHLGTYY 478

Scoring table: PAM 150
Gap 11

Searched: 116695 seqs, 37453910 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: p1r58
1:p1r1 2:p1r2 3:p1r3 4:p1r4

Statistics: Mean 48.417; Variance 93.758; scale 0.516

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	3126	90.2	478	1 S60754	transcription factor	0.00e+00
2	3032	87.5	486	1 TVHUEC	transferring protein	0.00e+00
3	1776	51.2	452	1 I37565	transferring protein	0.00e+00
4	1773	51.1	452	1 S17403	transferring protein	0.00e+00
5	1740	50.2	453	1 S49013	transferring protein	0.00e+00
6	1571	45.3	272	2 A54617	transferring protein	6.02e-302
7	1401	40.4	385	1 S29844	transferring protein	5.17e-265
8	745	22.5	173	2 A56646	transferring protein	1.65e-124
9	691	19.9	179	2 S51226	transferring protein	3.93e-113
10	519	15.0	103	2 S28823	transferring protein	1.57e-77
11	467	13.5	94	2 S28820	transferring protein	5.79e-67
12	458	13.2	268	2 S11224	transferring protein	3.80e-65
13	458	13.2	438	1 S11225	transferring protein	3.80e-65
14	456	13.2	440	1 A35875	transferring protein	9.60e-65
15	456	13.2	441	1 A53988	transferring protein	9.60e-65
16	456	13.2	441	1 TVHUEE	transferring protein	9.60e-65
17	453	13.1	472	1 B53236	transferring protein	3.86e-64
18	452	13.0	211	2 S24300	transferring protein	6.14e-64
19	451	13.0	440	2 I48291	transferring protein	9.76e-64
20	452	13.0	441	1 TVCHTE	transferring protein	6.14e-64
21	452	13.0	454	1 A48146	GA-binding protein a1	6.14e-64
22	452	13.0	454	1 A40858	GA-binding protein a1	6.14e-64
23	451	13.0	454	2 I38739	nuclear respiratory f	9.76e-64

24	452	13.0	464	1	S37616	transcription factor	6.14e-64
25	450	13.0	472	1	A53236	transcription factor	1.55e-63
26	452	13.0	485	1	TVCHET	transcription factor	6.14e-64
27	447	12.9	479	1	TVCHER	transcription factor	6.22e-63
28	448	12.9	488	1	TVFES	transcription factor	3.92e-63
29	445	12.8	468	1	TVHSE2	transcription factor	1.57e-62
30	445	12.8	469	1	TVHSE2	transcription factor	1.57e-62
31	440	12.7	110	2	A45938	transcription factor	1.59e-61
32	434	12.5	623	1	S33167	pointed protein, split	2.53e-60
33	434	12.5	718	1	S33168	pointed protein, split	2.53e-60
34	425	12.3	477	1	B46396	transcription factor	1.60e-58
35	425	12.3	477	1	I38893	transcription factor	1.60e-58
36	419	12.1	510	1	S43692	transcription factor	2.53e-57
37	404	11.7	342	2	A46396	ets-related protein 7	2.45e-54
38	407	11.7	462	1	S35534	adenovirus E1A enhanc	6.20e-55
39	407	11.7	555	1	S24061	transcription factor	6.20e-55
40	397	11.5	548	2	S59133	ETS2 repressor factor	5.99e-53
41	395	11.4	428	1	TVHUEK	transforming protein	1.49e-52
42	395	11.4	429	2	JC4965	elk1 protein - mouse	1.49e-52
43	394	11.4	430	2	I48755	MSAP1A - mouse	2.35e-52
44	393	11.3	250	2	A54308	PE-1 protein - human	3.71e-52
45	383	11.0	453	2	B42093	serum response factor	3.51e-50

ALIGNMENTS

RESULT 1
ENTRY S60754 #type complete
TITLE transcription factor erg - chicken
ORGANISM Gallus gallus #common_name chicken
DATE 27-Apr-1996 #sequence_revision 23-May-1997 #text_change 10-Jul-1998

ACCESSIONS
REFERENCE S60754
#authors Dhordain, P.; Dewitte, F.; Desbiens, X.; Stehelin, D.; Duterque-Coguiland, M.; Mech. Dev. (1995) 50:17-28
#journal Mesodermal expression of the chicken erg gene associated with precartilaginous condensation and cartilage differentiation.
#title

#cross-references M01D:95329425
#accession S60754
#status preliminary
#molecule_type mRNA
#residues 1-478 #label DHO
#cross-references EMBL:X77159; NID:g790439; PID:g790440
GENETICS

KEYWORDS
FEATURE 119-193
312-390
SUMMARY
#domain ets RII regulatory region homology #label ETS2\
#domain ets DNA-binding domain homology #label ETS
#length 478 #molecular-weight 53913 #checksum 547

Query Match 90.2%; Score 3126; DB 1; Length 478;
Best Local Similarity 92.1%; Pred. No. 0.00e+00;
Matches 441; Conservative 10; Mismatches 26; Indels 2; Gaps 2;

Db	1	MASTIKREALSVSEDSLFECAVGSPLAKTEMASSSSESGQTSKMPRVQOWLSOP	60
Qy	1	MASTIKREALSVSEDSLFECAVGSPLAKTEMASSSSESGQTSKMPRVQOWLSOP	60
Db	61	PARVITKECNPNQVNGSRNSPDDCSVAKGKMYSSSNMGNGSYAEKHIPPNNMTT	120
Qy	61	PARVITKECNPNQVNGSRNSPDDCSVAKGKMYSSSNMGNGSYAEKHIPPNNMTT	120
Db	121	NRRVIVPADPTLWSTDHVROWLEMAVKEYGIPVDVILLFONIDGELCKKTKDDFORLT	180
Qy	121	NRRVIVPADPTLWSTDHVROWLEMAVKEYGIPVDVILLFONIDGELCKKTKDDFORLT	180

```
Db      181    PSYNDILSHLHYLETPLPHLTSDVDALONSPRLMARNTGATFTFPNTPSYPEA   240
        |||||||
Oy      181    PSYNDILSHLHYLETPLPHLTSDVDALONSPRLMARNTGATFTFPNTPSYPEA   240
Db      241    TORITTRPDLPEQARRSAMTSHSHPOTSKATOPSSSTVTKETEDORQLDPYOILGPTSS   300
        |||||||
Oy      241    TORITTRPDLPEQARRSAMTSHSHPOTSKATOPSSSTVTKETEDORQLDPYOILGPTSS   300
Db      301    RLANPGSGOIOLMOFLLELLSDSSNSNCITWEGNGEKKATDPEVARRMGERKSKRNAN   360
        |||||||
Oy      301    RLANPGSGOIOLMOFLLELLSDSSNSNCITWEGNGEKKATDPEVARRMGERKSKRNAN   360
Db      361    YKLRSALYYDYDKNIMTWKHGRKYAVKPEDFHGAQAOLQHPPRESSMYKYP-SOLPYMSS   419
        |||||||
Oy      361    YKLRSALYYDYDKNIMTWKHGRKYAVKPEDFHGAQAOLQHPPRESSMYKYP-SOLPYMSS   419
Db      420    YAHAPQKNMFVAHPRALPVTSSEFFAAPRPYNNSPGIGIYPTRTLPAAMPSHLGTYY   478
        |||||||
Oy      420    YAHAPQKNMFVAHPRALPVTSSEFFAAPRPYNNSPGIGIYPTRTLPAAMPSHLGTYY   478
ENTRY   2          TVHUEG #type complete
TITLE   transforming protein erg-3 - human
CONTAINS #normal_name Homo sapiens #common_name man
ORGANISM 31-Mar-1989 #sequence_revision 30-May-1997 #text_change
DATE     31-Oct-1997
ACCESSIONS A94294; A94178; I58410; A28041; A29515
REFERENCE A94294
#authors Rao, V.N.; Papas, T.S.; Shyam, E.; Reddy, P.
#journal Science (1987) 237:635-639
#title erg, a human ets-related gene on chromosome 21: alternative
#cross-references MUID:87263429 splicing, polyadenylation, and translation.
#accession A94294
##molecule_type mRNA
##residues 1-231,256-486 ##label REA
#cross-references GB:M17254; NID:g182186; PID:g182187
REFERENCE A94178
#authors Reddy, E.S.P.; Rao, V.N.; Papas, T.S.
#journal Proc. Natl. Acad. Sci. U.S.A. (1987) 84:6131-6135
#title The erg gene: a human gene related to the ets oncogene.
#cross-references MUID:87317608
#accession A94178
##molecule_type mRNA
##residues 100-231,256-486 ##label REB
#cross-references GB:M1535; NID:g182182; PID:g182185
REFERENCE I58410
#authors Prasad, D.D.; Rao, V.N.; Lee, L.; Reddy, E.S.
#journal Oncogene (1994) 9:669-673
#title Differentially spliced erg-3 product functions as a
transcriptional activator.
#cross-references MUID:94119611
#accession I58410
##status preliminary: translated from GB/EMBL/DDBU
##molecule_type mRNA
##residues 230-259 ##label RES
#cross-references GB:S68130; NID:g544960; PID:g544961
GENETICS GDB:ERG
#gene GDB:ERG
#cross-references GDB:119884; OMIM:165080
CLASSIFICATION #map_position 21q22.2-21q22.2
#superfamily transcription factor erg; ets DNA-binding domain
homology; ets RII regulatory region homology
KEYWORDS alternative splicing; DNA binding; nucleus; proto-oncogene;
transcription factor; transforming protein
FEATURE 1-231,256-486 #product transforming protein erg-2 #status predicted
100-231,256-486 #product transforming protein erg-1 #status predicted
126-200 #domain ets RII regulatory region homology #label ETS\
```

Query Match	87.5%	Score 3032	DB 1	Length 486
320-398	#domain ets DNA-binding domain homology	#label ETS		
SUMMARY	#length 486	#molecular-weight 54608	#checksum 9577	
Best Local Similarity	87.5%	Pred. No. 0.00e+00		
Matches 419	Conservative 28	Mismatches 29	Indels 3	Gaps 3

Query Match	87.5%	Score 3032	DB 1	Length 486
Db	9	AAHKELALSVSEDSQSLFECAATGTPHLATTEMTASSSSDYGTSTKMSRRVYQOOWLSQPP	68	
Qy	2	ASTKEALSVSEDSQSLFECAATGTPHLATTEMTASSSSDYGTSTKMSRRVYQOOWLSQPP	61	
Db	69	ARVITKECPNOVNGSRNSPDCSCVAKGKKVAGSDPTGNNYGSYMEKEKHPMPNMTN	128	
Qy	62	ARVITKECPNOVNGSRNSPDCSCVAKGKKVAGSDPTGNNYGSYMEKEKHPMPNMTN	121	
Db	129	ERRVIVPADPLTMSDHWQMLEMAVKEYGLPDVNIILLFQNIIDGKELCKMTKDDFORLTP	188	
Qy	122	ERRVIVPADPLTMSDHWQMLEMAVKEYGLPDVNIILLFQNIIDGKELCKMTKDDFORLTP	181	
Db	189	SYNADILLSHVIREPLPHLITSDVYDALONSPLMHAARTGCAATFPNTSYPEAT	248	
Qy	182	SYNADILLSHVIREPLPHLITSDVYDALONSPLMHAARTGCAATFPNTSYPEAT	241	
Db	249	QRTTRPDLPEPPRSRGAATGSHGPPRQSKAOPSSTYPTKEDORPOLDYQIILGPS	308	
Qy	242	QRTTRPDLPEPPRSRGAATGSHGPPRQSKAOPSSTYPTKEDORPOLDYQIILGPS	300	
Db	309	RLANPGSGQIQLOMQLLELLSDSSNSCITWETNGEFTMDPDEVARMERKSKPMN	368	
Qy	301	RLANPGSGQIQLOMQLLELLSDSSNSCITWETNGEFTMDPDEVARMERKSKPMN	360	
Db	369	YDKLSRLRRYYDKNIMTKYHGKRRYAKDFHGLAOLAPHPRESSLYKYP-SDLPIWGS	427	
Qy	361	YDKLSRLRRYYDKNIMTKYHGKRRYAKDFHGLAOLAPHPRESSLYKYP-SDLPIWGS	419	
Db	428	YHAPOKMNVAAPRPALPYTSSSFEFAANPWNPSPTGIYNTPLRSHHPSHLGYTY	486	
Qy	420	PHAPQKMNVAAPRPALPYTSSSFEFAANPWNPSPTGIYNTPLRSHHPSHLGYTY	478	

RESULT	ENTRY	3
TITLE	137565	#type complete
ALTERNATE_NAMES	transforming protein fll, long splice form - human	
ORGANISM	friend leukemia integration protein 1; transcription factor	
DATE	ERGB	
ACCESSIONS	#formal_name Homo sapiens #common_name man	
REFERENCE	04-Oct-1996 #sequence_revision 30-May-1997 #text_Change	
	20-Mar-1998	
	137565; S29843; S35506; A49000; A49015; I54170	
	I37378	
	Delattre, O.; Zucman, J.; Plougastel, B.; Desmaze, C.; Melot, T.; Peter, M.; Kovar, H.; Joubert, I.; de Jong, P.; Rouleau, G.; Aurias, A.; Thomas, G.	
	Nature (1992) 359:165-165	
	Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours.	
	#cross-references MUID:92396239	
	#accession I37565	
	#status preliminary; translated from GB/EMBL/DBJ	
	##molecule_type mRNA	
	##residues 1-452 ##label RES	
	##cross-references EMBL:X67001; NID:g32529; PID:g32530	
	REFERENCE S29843	
	#authors	
	Homas, R.; May, W.; Denny, C.; Raskind, W.; Moore, J.; Makl, R.A.; Beck, E.; Klemenz, M.J.	
	Biochim. Biophys. Acta (1993) 1172:155-158	
	Human FLI-1 localizes to chromosome 11Q24 and has an aberrant transcript in neuroepithelioma.	
	#cross-references MUID:9316799	
	#accession S29843	
	##molecule_type mRNA	
	##residues 1-68; 'V', '70-129', 'A', '131-132', 'V', '134-322', 'Q', '324-425', '427-452' ##label HNO1	

[illegible]

```

0y 1 MASTIKELVSVSDQSLFECACGSP -HLAKTEMTASSSSEYQITKMSPRPRQDMLSQ 59
Db 61 P-VARNVKNREYD--HMNGSRESPVDCSVSKCNKLVGCGEANPNYNSYMDERKNPPPM 117
0y 60 PPAVTTIKMECNPNQVNGSRNSPDDCSVAAGKMKVSSSDVNGVNGSYMEERK -IIPPMM 118
Db 118 TTNERARYVPADPFLTMQOEHYROWLEWAIKEYGMEIDTSPFONMNGKELCKMKNEDFLR 177
0y 119 TTNERARYVPADPFLMNSTIDHYROWLEWAAVEYGLPVDVILFQNDIKELCKMTKDOR 178
Db 178 ATSAVNTFVLLSHLSYLRESSL--LAYNTTSH-TDQSSRL-NYKE--D----P-S-Y- 222
0y 179 LTPSNADILLSHHYLRETPFLPLTSDVDKALONSPRLMHARNMGATFTFPNTSVYP 238
Db 223 DSV-R---R-G-ANNMNSG-LNKS-PLLG-OSQ-----TMGNTORQORPQYOLGPT 269
0y 239 EATQITTRPDLPEQARRSMTSHSHPTOSKATQPBSSVTPEQORQOLDPYOLGPT 298
Db 270 SSRLANPESGQIQLOMQLLEHLLSDSAASCITGEGTNGEKKMDPEVARRMGERKSKEN 329
299 SSRLANPESGQIQLOMQLLEHLLSDSSNSCITGEGTNGEKKMDPEVARRMGERKSKEN 358
Db 330 MNVOKLSAALAYYDDKIMTKVHGKRAAYKFDHGIQAQALPPTTSMYKTP-SDISYM 388
0y 359 MNVOKLSAALAYYDDKIMTKVHPPRESS-WYKYPSPDLPYSSSYHGKRAYKAKDFHGIAHA 417
Db 389 PSYHAHQKLVFVPSHSSSMPTSSSFEGCAASQWYTPSTGATIPNRSVRHNTHTVPSHL 448
0y 418 LQPHAHQKMNFAVRHPRALPYTSSSFAPARNPYMSPGIGIYR-T-RLDAAHPSHL 474
Db 449 GSXY 452
0y 475 GTYY 478

```

```

RESULT      5
ENTRY
TITLE       S49013 #type complete
ALTERNATE_NAMES  transforming protein fl1 - African claved frog
              Friend leukemia integration protein 1; transcription factor
              ERFB
ORGANISM    #formal_name Xenopus laevis #common_name African claved frog
DATE        07-May-1995 #sequence_revision 23-May-1997 #text_change
            05-Sep-1997
ACCESSIONS  S49013
REFERENCE   S49013
AUTHORS     Meyer, D.; Wolff, C.M.; Stiegler, P.; Senan, F.; Befort, N.;
            Befort, J.C.; Remy, P.
            Mech. Dev. (1993) 44:109-121
            xl-fl1, the Xenopus homologue of the fl1-1 gene, is expressed
            during embryogenesis in a restricted pattern evocative of
            neural crest cell distribution.
            S49013
            preliminary
#accession  ##status
#molecule_type mRNA
##residues  1-453 ##label MEY
#cross-references EMBL:X66979; NID:g505486; PID:g505487
GENETICS
#gene       fl1
CLASSIFICATION  #superfamily transcription factor eeg; ets DNA-binding domain
                 homology; ets RII regulatory region homology
KEYWORDS       DNA binding; nucleus; proto-oncogene; transcription factor;
                 transforming protein
FEATURE
117-191       #domain ets RII regulatory region homology #label ETS2\
284-362       #domain ets DNA-binding domain homology #label ETS
SUMMARY       #length 453 #molecular_weight 51015 #checksum 774
Query Match   50.2%; Score 1740; DB 1; Length 453;
Best Local Similarity 56.4%; Pred. No. 0.00e+00;
Matches 273; Conservative 91; Mismatches 83; Indels 37; Gaps 25;
db           1 MDGTTKALSVSDDSGLFDSAAGASHLSKADMTASANDPYCQPHKINIRPPQQMINDQ 60
| : :::::::::::::::::::: |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

```

0Y      1 MASTIKELSVSEDSQSLFEBCATG-SPHLAKTEMTASSSESYGOTSMSRVPVQOOWLSQ 59
Db      61 P-MRVNIREXE--HMNGSRSPVDCSINKCSKILGSEGNAMTY-TYDEKNGPEPPN 116
0Y      60 PPARVITMECNPNQOVNRSRNSPDDCSVAGGKVVSSDVMGNYSGYMEKH-IPPNN 118
Db      117 TTNERRVVPPADPALMSODHROWLEWAIREYGEVLEJDCSLFONIDGKELCKSKKEPDR 176
0Y      119 TTNERRVVPPADPLMTSDHROWLEWAIREYGLPVDVILLFONIDGKELCKMTKODFOR 178
Db      177 STSIINTEVLLSHLNYLSDS-SSLGYNTOANT-DOSSRLT-AKE--D---P-S-Y- 222
0Y      179 LTPSNADILLSHLYLRETLPLHLTSDVDYKALONSPRLMHRNNTGATFIPRNTSVY 238
Db      223 EAV-R---RSG--WGNMSNP-VYKSPPM-G-GTQ-VNKN-S-GDOQRSDPDTYOLIGP 270
0Y      239 EATQITTRPDLPEYQARSAMTSHSPHOTOSKARPOSSVTPXTEODRPDLDTYOLIGP 298
Db      271 SSRLANPESGOIQLOMQLFLELLSDSSNASCSITWEGTNGEERKMTDPDEVARMEERSKPN 330
0Y      299 SSRLANPESGOIQLOMQLFLELLSDSSNSNCITWGTNGEERKMTDPDEVARMEERSKPN 358
Db      331 MNYDLASALXYTUDKSMTYVHBKRRAYKEDFNGIALQAPRPDTSMKY-P-SEFSYA 389
0Y      359 MNYDLASALXYTUDKMTKMTVHPRESS-WKXPYSDLPYSSSYHGKRYAKKFFEHGIAQA 417
Db      390 PSYSHQOQVNFVPSHPSPMPTSSGFGATSPWNSPANISYIPNVPVPHPTVHQSIL 449
0Y      418 LQPHNHPQKNMFVAPHRPALVYSSSFAPANRPYWNTPSGITPNT---RLPAAHNPSHL 474
Db      450 GGFY 453
0Y      475 GYY 478

```

```

RESULT      6
ENTRY
TITLE
ALTERNATE_NAMES
ORGANISM
DATE

ACCESSIONS
REFERENCE
#authors
#journal
#title
#cross-references MUID:94019387
#accession      A54617
#status          preliminary
#molecule_type mRNA
#residues        1-272
##experimental_source pre-B-cell line 22D6
##note           sequence extracted from NCBI backbone (NCBIN:138523,
                    NCBI:P138524)

CLASSIFICATION
KEYWORDS
FEATURE
140-218
SUMMARY
#domain ets DNA-binding domain homology #label ETS
#length 272 #checksum 3754

Query Match      45.3%; Score 1571; DB 2; Length 272;
Best Local Similarity 82.4%; Pired. No. 6.02e-302;
Matches 225; Conservative 17; Mismatches 28; Indels 3; Gaps 3.

Db      1 DDFORLTPSYNDILLSHYLREPLPLHLSDDVVKALONSRMLHARTGTGAFFPN 60
        |DFORLTPSYNDILLSHYLREPLPLHLSDDVVKALONSRMLHARTGTGAFFPN 233
QY      174 DDFORLTPSYNDILLSHYLREPLPLHLSDDVVKALONSRMLHARTGTGAFFPN 233

Db      61 TSVYEATQRIITRPLDPEPRRSAMTGHSHLTPOSKAQPSPAVPKTEDPQPOLDPY 120

```

Oy	234	TSVPEATGRTITRPDLPEQARSAWTSHTP-OSKATQSSSTVPTKPEDQRPOLDPY	292
Db	121	QILGPTSSRLANPGSGQIQIQWQFLLELLSDSSNSNCITWGTNGERKMTDPDEYARWGE	180
Oy	293	QILGPTSSRLANPGSGQIQIQWQFLLELLSDSSNSNCITWGTNGERKMTDPDEYARWGE	352
Db	181	RKSKNNMNDKLSRALRYYYDKNIMKRVGKRAYVDFEFGIALQAPHPPESSLYKYP-	239
Oy	333	RKSKNNMNDKLSRALRYYYDKNIMKRVGKRAYVDFEFGIALQAPHPPESSLYKYP-	411
Db	240	SDLPYMGSYHAHPKMKNEVSPHPALPYTSSSF	272
Oy	412	HGIALQAPHPKMKNEVSPHPALPYTSSSF	444
RESULT	7		
ENTRY	S29844	#type complete	
TITLE	transforming protein flt1, short splice form - human		
ALTERNATE_NAMES	Flt1 leukemia integration protein 1; transcription factor ERB		
ORGANISM	Human		
DATE	02-Sep-1997		
ACCESSIONS	S29844		
REFERENCE	S29843		
#authors	Hromas, R.; May, W.; Denny, C.; Raskind, W.; Moore, J.; Makl, R.A.; Beck, E.; Klemenz, M.J.		
#journal	Biochim. Biophys. Acta (1993) 1172:155-158		
#title	Human Flt-1 localizes to chromosome 11Q24 and has an aberrant transcript in neuroepithelioma.		
#cross-references	MUID:93176799		
#accession	S29844		
#status	preliminary; nucleic acid sequence not shown; translation not shown		
GENETICS			
#gene	GDB:FLI1		
#residues	1-385		
#cross-references	EMBL:M93255; NID:g18265; PID:g182651		
#note	the nucleotide sequence was submitted to the EMBL data library, May 1992		
CLASSIFICATION			
#map_position	11q24.1-11q24.3		
KEYWORDS	transcription factor; transforming protein		
FEATURE			
52-126	#domain ets R11 regulatory region		
217-295	#domain ets DNA-binding domain		
SUMMARY			
Query Match	40.4%	Score 1401; DB 1; Length 385;	
Best Local Similarity	56.4%	Pred. No. 5.17e-265;	
Matches	229; Conservative	69; Mismatches 73; Indels 35; Gaps 21;	
Db	10	ASRPDSCVSCKTVGGESPMYNSYNDKNGKPPNNMTNERRVYVPADATLVYQ	69
Oy	78	SKNSPDSCVAAGGKRVSSDNGMYNWKERH-IPPNMTNERRVYVPADATLVYQ	136
Db	70	EHHROWLENAIKKEYSLAMEIDTSFFQOMDGEKELCKMKNEDELRATTVLTNTEVILSHLSYLR	129
Oy	137	DHVRQLENAVAVEYGLPDVDILLFQINIDGKELCKMKRDKDQRLPLPSVNAIDILSHLSYLR	196
Db	130	ESSL-LAYNTTSHT-DQSSRL-SYKE--D---P--S-Y-DSV-R--R-G-AMGNM	168
Oy	197	ETPLPLTSDVDKALQNSRLMHAHNTGATFIFPNTSYVPEATGRTITRPDLPEQAR	256
Db	169	NSG-LINKS-PPLG-GAO---TISKTEORQPPDQIQIILPTSSRLANPGSGQIQIQWQFL	221
Oy	257	NSAMTSHSPQSKATQPPSSSTVPTKPEDQRPOLDPYQILPTSSRLANPGSGQIQIQWQFL	316

Db	222	LELLSDANASCITMGCTGEEFMTDPDEVARWGORSKSPMANTDKSRLRYTODKI	281
OY	317	LELLSDSSNSNCITMGCTGEEFKMTDPDEVARWGERKSPMANTDKSRLRYTODKI	376
Db	282	MTKVGKRYAYKDFHGIAOALOPHPTESSMYKYP-SDISYWP5YAHQORVNEVPHP5	340
OY	377	MTKVHPRESS-MTKYPSDDLPMSSSYGKRYAKKPFDFHGIAOALOPHAPRQKMFVAPHP	435
Db	341	SMPYTSSEFFGAASQYWT5-TGGITPYNVPNPHPTVHP5HLG5TY	385
OY	436	ALPVTSSSEFFAAPNPYWSNPTG5TYENT--RLPAAHMPSHLG5TY	478
RESULT	8		
ENTRY		A56646	#type fragment
TITLE		transcription factor erg/fil-1 homolog - sea urchin (Lytechinus variegatus) (fragment)	
ORGANISM		formal_name Lytechinus variegatus #common_name variegated urchin	
DATE		03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 08-Sep-1997	
ACCESSIONS		A56646	
REFERENCE		A56646	
#authors		Qi, S.; Chen, Z.O.; Papas, T.S.; Lautenberger, J.A.	
#journal		DNA Seq. (1992) 3:127-130	
#title		The sea urchin erg homolog defines a highly conserved erg-specific domain.	
#accession		A56646	
#status		preliminary	
#molecule_type		DNA	
#residues		1-173 ##label Q1A	
CLASSIFICATION		##cross-references GB:M1067; NID:g161310; PID:g161311	
KEYWORDS		superfamily transcription factor erg; ets DNA-binding domain	
FEATURE		homology; ets RII regulatory region homology	
SUMMARY		DNA binding; nucleus; transcription factor	
		#domain ets DNA-binding domain homology #label ETS	
		#length 173 #checksum 5090	
Query Match			
Best Local Similarity	21.5%:	Score 745; DB 2; Length 173;	
Matches	110; Conservative	30; Mismatches 26; Indels 12; Gaps 11;	
Db	1	SGGIQLMOFLLELLSDSSNANCITMGTEGEEFKMTDPDEVARWGERKSKPMNYDKLSR	60
OY	307	SGGIQLMOFLLELLSDSSNSNCITMGTEGEEFKMTDPDEVARWGERKSKPMNYDKLSR	366
Db	61	ALRYYDKNMTVHGKRRR-1KFPDAGLAQMOPVQADPSKRYQ-SDLR-YLPGR-HP	116
OY	367	ALRYYDKNMTVHPPESSMYKYP-SDLP-TMSSYHGKRYAYKDFHGIAOALOPHAPR	424
Db	117	TKLWVGT-LIN-PSINASLFSHSSYSSPGANINYP5GHTPHHAAHMSHLG5TY	172
OY	425	QKMFVAPHPALPVTSSEFFAAPNPYWSNPTG5TYENT--P-AAHMP5HLG5TY	478
RESULT	9		
ENTRY		S51226	#type fragment
TITLE		transcription factor erg/fil-1 homolog - polychaete (Nereis diversicolor) (fragment)	
ORGANISM		#formal_name Nereis diversicolor #common_name sandworm	
DATE		15-Jul-1995 #sequence_revision 23-May-1997 #text_change 30-May-1997	
ACCESSIONS		S51226	
REFERENCE		S51226	
#authors		Lelievre-Chocteau, A.; Landet, V.; Flourens, A.; Begue, A.; Lepince, D.; Fontaine, F.	
#journal		FEBS Lett. (1994) 354:62-66	
#title		Identification of two ets related genes in a marine worm, the polychaete annelid Nereis diversicolor.	
#accession		S51226	
#status		preliminary	
#molecule_type		DNA	

	##residues	1-179 ##label LEL	
CLASSIFICATION	#superfamily transcription factor erg; ets DNA-binding domain		
KEYWORDS	homology; ets RII regulatory region homology		
FEATURE	DNA binding; nucleus; transcription factor		
7-85			
SUMMARY	#domain ets DNA-binding domain homology #label ETS		
	#length 179 #checksum 4104		
Query Match	19.9% Score 691; DB 2; Length 179;		
Best Local Similarity	57.5% Pred. No. 3,93e-113;		
Matches	104; Conservative 36; Mismatches 28; Indels 13; Gaps 12;		
Db	1 GSGQIQLOMFLLELLSDANSIIITWGTGCEKLVDPETARRRGSKSPNNNYDKLS 60		
Oy	306 GSGQIQLOMFLLELLSDSNSNCITWGTGCEKMDPDEVARRRGRKSKPPNNYDKLS 365		
Db	61 RALRYDYDNIMTKVHGKRKYA-YKPFAGIAQMOPSTTDPAAYKKYO-QDLMS-GYH-H 116		
Oy	366 RALRYDYDNIMTKVHPPESSMYKYP-SDLP-YMSSYHGKRKYAKPFDFHQAQLQPHAH 423		
Db	117 TSKLNLMAAHAP-MASSASGFPPPPAPYWSLYGSNLYPIINSHAMSHHFGCHSSHLCGY 175		
Oy	424 POKMNFVAHPHALPYTSSFFAAPNPYWSPTGG-TYPN-T-R-LP-AAHMPSHDGT 477		
Db	176 Y 176		
Oy	478 Y 478		
RESULT	10		
ENTRY	S28823	#type fragment	
TITLE	transcription factor cts-21C - fruit fly (Drosophila melanogaster) (fragment)		
ALTERNATE_NAMES	transforming protein cts-6		
ORGANISM	#formal_name Drosophila melanogaster		
DATE	31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 23-May-1997		
ACCESSIONS	S28823		
REFERENCE	S28819		
#authors	Chen, T.; Bunting, M.; Karim, F.D.; Thummel, C.S.		
#journal	Dev. Biol. (1992) 151:176-191		
#title	Isolation and characterization of five Drosophila genes that encode an cts-related DNA binding domain.		
#accession	S28823		
#molecule_type	DNA		
#residues	1-103 ##label CHE		
#cross-references	EMBL:M88475		
GENETICS			
#gene	Erg21C		
#coss-references	FlyBase:Fggn0005660		
#intons	64/2		
CLASSIFICATION	#superfamily transcription factor erg; ets DNA-binding domain		
KEYWORDS	homology; cts RII regulatory region homology		
FEATURE	DNA binding; nucleus; transcription factor		
8-86			
SUMMARY	#domain ets DNA-binding domain homology #label ETS		
	#length 103 #checksum 5825		
Query Match	15.0% Score 519; DB 2; Length 103;		
Best Local Similarity	77.6% Pred. No. 1.57e-77;		
Matches	66; Conservative 11; Mismatches 7; Indels 1; Gaps 1;		
Db	2 SGGQIQLOMFLLELLSDNSNANAISWEGSGEFLIDPDVARRRGRKAKPPNNYDKLS 61		
Oy	306 GSGQIQLOMFLLELLSDSNSNCITWGTGCEKMDPDEVARRRGRKSKPPNNYDKLS 365		
Db	62 RALRYDYDNIMTKVHGKRKYA-YKF 85		
Oy	366 RALRYDYDNIMTKVHPPESSMYKY 390		
RESULT	11		
ENTRY	S28820	#type fragment	

TITLE		transcription factor ets-65A - fruit fly (<i>Drosophila melanogaster</i>) (fragment)
ALTERNATE_NAMES		transforming protein ets-3
ORGANISM		formal_name Drosophila melanogaster
DATE		31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 23-May-1997
ACCESSIONS		S28820
REFERENCE		S28819
#authors		Chen, T.; Bunting, M.; Karin, F.D.; Thummel, C.S.
#journal		Dev. Biol. (1992) 151:176-191
#title		Isolation and characterization of five <i>Drosophila</i> genes that encode an ets-related DNA binding domain.
#accession		S28820
##molecule_type	DNA	
##residues	1-94	##label CHE
##cross-references	EMBL:M88473	
GENETICS		
#gene	Ets65A	##cross-references FlyBase:FBN0005658
CLASSIFICATION		#superfamily transcription factor crg; ets DNA-binding domain homology; ets RII regulatory region homology
KEYWORDS		DNA binding; nucleus; transcription factor
FEATURE		
37-94		#domain ets DNA-binding domain homology (fragment)
SUMMARY		#label ETS
	#length 94	#checksum 1156
Query Match	13.5%	Score 467; DB 2; Length 94;
Best Local Similarity	93.8%;	Pred. No. 5,79c-67;
Matches	61; Conservative	4; Mismatches 0; Indels 0; Gaps 0;
DB	30 PGSGIOLMOLFLELSDSNNSACTWEGTNGEFKTPDDEVARRWGGRSKSPNNYDKL	89
OY	305 PGSGIOLMOLFLELSDSNSNCITWGTNEFKTPDDEVARRWGGRSKSPNNYDKL	364
DB	90 SRALR 94	
OY	365 SRALR 369	
RESULT	12	
ENTRY	S11224	#type fragment
TITLE		transcription factor ets-1b - African clawed frog (fragment)
ORGANISM		formal_name Xenopus laevis #common_name African clawed frog
DATE		21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 08-Sep-1997
ACCESSIONS		S11224
REFERENCE		S11224
#authors		Stiegler, P.; Wolff, C.M.; Baltzinger, M.; Hitzelin, J.; Senan, F.; Meyer, D.; Ghysdael, J.; Stohelin, D.; Befort, N.; Remy, P.
#journal		Nucleic Acids Res. (1990) 18:5298
#title		Characterization of <i>Xenopus laevis</i> cDNA clones of the c-ets-1 proto-oncogene.
#cross-references	MOTID:90384849	
#accession	S11224	
##molecule_type	mRNA	
##residues	1-268	##label STI
##cross-references	EMBL:X52691; NID:g64616; PID:g64617	
CLASSIFICATION		#superfamily transcription factor ets; ets DNA-binding domain homology; ets RII regulatory region homology
KEYWORDS		DNA binding; nucleus; phosphoprotein; transcription factor
FEATURE		
164-242		#domain ets DNA-binding domain homology #label ETS
SUMMARY		#length 268 #checksum 3247
Query Match	13.2%	Score 458; DB 2; Length 268;
Best Local Similarity	51.7%;	Pred. No. 3,80c-65;
Matches	60; Conservative	25; Mismatches 30; Indels 1; Gaps 1;
DB	137 DRALNDKPYPVPAALAGCTGCGPQLMOLFLELTDKSCGSFISWTGDCMERFLSPD	196
OY	285 GRPLDPQILGPSSSLANPGSQIOLMOLFLELSDSNSNCITWGTNGEFTMDPD	344

#title	Sequence-specific DNA binding of the proto-oncoprotein ets-1 defines a transcriptional activator sequence within the long terminal repeat of the Moloney murine sarcoma virus.
#cross-references	M0ID:90299137
#accession	A35875
##molecule-type	mRNA
##residues	170-440 #label GUN
##cross-references	EMBL:X53953
GENETICS	Ets1
#gene	map_position 9 15.0
CLASSIFICATION	superfamily transcription factor ets; ets DNA-binding domain homology; ets RII regulatory region homology
KEYWORDS	alternative splicing; DNA binding; nucleus; phosphoprotein; proto-oncogene
FEATURE	
57-130	#domain ets RII regulatory region homology #label ETS\
97-130	#region helix-loop-helix #status predicted\
337-415	#domain ets DNA-binding domain homology #label ETS\
377-383	#region nuclear location signal
SUMMARY	#length 440 #molecular-weight 50201 #checksum 2673
Query Match	13.2%; Score 456; DB 1; Length 440;
Best Local Similarity	51.7%; Pred. No. 9,60e-65;
Matches	60; Conservative 25; Mismatches 30; Indels 1; Gaps 1;
Db	310 DRADLNKRPVIPAALAGYTGSGPIQMOFTLELITDKSCSFTISWGDGNEFKLSPD 369
Oy	285 QRFOLDPYOILGTGSRLNPGSGIQLOFLELLSDSSNSNCITWEGTNEEFMTDPD 344
Db	370 EVARFGCKRKNPKMYEKLRSGLRRYYDNKIHT-AGRKYRVFVCDDLSLG 424
Oy	345 EVARFGCKRKNPKMYDKLSRALRYDKNITVTVPPESSMYRYPDDLPMSSY 400
RESULT	15
ENTRY	A53988
TITLE	#type complete
ALTERNATE_NAMES	transcription factor ets-1, splice form a - rat
ORGANISM	transcription factor p54; transforming protein ets-1a; transforming protein Tpl-1
DATE	#formal_name Rattus norvegicus #common_name Norway rat 09-Oct-1994 #sequence_revision 23-May-1997 #text_change 23-May-1997
ACCESSIONS	A53988
REFERENCE	A53988
#authors	Bellacosa, A.; Datta, K.; Bear, S.E.; Patrioticis, C.; Lazo, P.A.; Copeland, N.G.; Jenkins, N.A.; Tischlis, P.N.
#journal	J Virol. (1994) 68:2320-2330
#title	Effects of provirus integration in the Tpl-1/Ets-1 locus in Moloney murine leukemia virus-induced rat T-cell lymphomas levels of expression, polyadenylation, transcriptional initiation, and differential splicing of the Ets-1 mRNA.
#cross-references	MID:94187072
#accession	A53988
##status	preliminary
##molecule-type	mRNA
##residues	1-441 #label BEL
##cross-references	GB:L20681; NID:g404781; PID:g404782
GENETICS	Ets-1
CLASSIFICATION	superfamily transcription factor ets; ets DNA-binding domain homology; ets RII regulatory region homology
KEYWORDS	alternative splicing; DNA binding; nucleus; phosphoprotein; proto-oncogene
FEATURE	
57-130	#domain ets RII regulatory region homology #label ETS\
97-130	#region helix-loop-helix #status predicted\
337-415	#domain ets DNA-binding domain homology #label ETS\
377-383	#region nuclear location signal
SUMMARY	#length 441 #molecular-weight 50422 #checksum 6268
Query Match	13.2%; Score 456; DB 1; Length 441;
Best Local Similarity	51.7%; Pred. No. 9,60e-65;

	Matches	60;	Conservative	25;	Mismatches	30;	Indels	1;	Gaps	1;
Db	310	DRADLNKDKPVIPAAALAGYTGSGPIQLWQFLLELLTDKSCQSFISWTGDNWFEKLSDPD	369							
Oy	285	ORPOLDPYQILGPTSSRLANPGSGOIQLMQFLLELLSDSSNSNCITWECTNGEFKMTDPD	344							
Db	370	EYARRWGKRKNKPKMNYEKLRSGLRYYYDKNIHKT-AGKRYVRFCVCDLQSLGy	424							
Oy	345	EYARRWGERKSKPMNMNDKLSRALRYYYDKNIMTKVHPPESSMYKYPDDLPMSSy	400							

Search completed: Wed Apr 7 09:44:31 1999
Job time : 78 secs.

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

W E S E R E H (TM)

Search: protein - protein database search, using Smith-Waterman algorithm
Run on: Wed Apr 7 09:44:50 1999; Maspar time 13.89 Seconds
Molecular output not generated. 923.482 Million cell updates/sec

Title: >US-08-878-177-4
Description: (1-478) from US08878177.pep
Perfect Score: 3467
Sequence: 1 MASTIKELSVSEDSLE.....IYPNRLPAHMPHSLGTY 478

Scoring table: PAM 150
Gap 11

Searched: 74019 seqs, 26840295 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot35
1:swissprot

Statistics: Mean 49.703; Variance 83.064; scale 0.598

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	2608	75.2	462	1	ERG_HUMAN TRANSFORMING PROTEIN E	0.00e+00
2	1776	51.2	452	1	FLI1_HUMAN FLI-1 ONCOGENE (ERGB T	0.00e+00
3	1773	51.1	452	1	FLI1_MOUSE RETROVIRAL INTEGRATION	0.00e+00
4	1740	50.2	453	1	FLI1_XENLA RETROVIRAL INTEGRATION	0.00e+00
5	1571	45.3	272	1	ERG_MOUSE TRANSFORMING PROTEIN E	0.00e+00
6	745	21.5	173	1	ERG_LYTVA ERG PROTEIN HOMOLOG (F	1.26e-14
7	519	15.0	102	1	ETS6_DROME DNA-BINDING PROTEIN D-	1.46e-75
8	459	13.2	64	1	ETS3_DROME DNA-BINDING PROTEIN D-	2.49e-75
9	458	13.2	268	1	ET1B_XENLA C-ETS-1B PROTEIN (XEL-	2.49e-75
10	458	13.2	438	1	ET1A_XENLA C-ETS-1A PROTEIN.	7.28e-75
11	456	13.2	440	1	ETS1_MOUSE C-ETS-1 PROTEIN (P54).	7.28e-75
12	456	13.2	441	1	ETS1_RAT C-ETS-1 PROTEIN (P54).	7.28e-75
13	456	13.2	441	1	ETS1_HUMAN C-ETS-1 PROTEIN (P54).	7.28e-75
14	453	13.1	472	1	ETS2A_XENLA C-ETS-2A PROTEIN.	3.63e-74
15	452	13.0	441	1	ETS2A_CHICK TRANSFORMING PROTEIN P	6.19e-74
16	452	13.0	454	1	GABA_HUMAN GA BINDING PROTEIN ALP	6.19e-74
17	452	13.0	454	1	GABA_MOUSE GA BINDING PROTEIN ALP	6.19e-74
18	452	13.0	464	1	ELG_DROME DNA-BINDING PROTEIN D-	6.19e-74
19	450	13.0	472	1	ETS2B_XENLA C-ETS-2B PROTEIN.	1.81e-73
20	452	13.0	485	1	ETSB_CHICK TRANSFORMING PROTEIN P	6.19e-74
21	447	12.9	479	1	ETS2_CHICK C-ETS-2 PROTEIN.	8.97e-73
22	448	12.9	669	1	MYBE_AVILE P135-GAG-MYB-ETS TRANS	5.26e-73
23	445	12.8	468	1	ETS2_MOUSE C-ETS-2 PROTEIN.	2.61e-72

ALIGNMENTS

RESULT	1	STANDARD:	PRT:	462 AA.
ID	ERG_HUMAN			
AC	P11308;			
DT	01-JUL-1989 (REL. 11, CREATED)			
DT	01-JUL-1989 (REL. 11, LAST SEQUENCE UPDATE)			
DT	01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)			
DE	TRANSFORMING PROTEIN ERG (ERG-2 / ERG-1).			
GN	ERG.			
OS	HOMO SAPIENS (HUMAN).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;			
CC	EUTHERIA; PRIMATES.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
EX	MEDLINE; 87263429.			
RA	RAO V.N., PAPAS T.S., SHYAM E., REDDY P.;			
RL	SCIENCE 237:635-639(1987).			
RN	[2]			
RP	SEQUENCE OF 100-462 FROM N.A.			
EX	MEDLINE; 87317608.			
RA	REDDY E.S.P., RAO V.N., PAPAS T.S.;			
RL	PROC. NATL. ACAD. SCI. U.S.A. 84:6131-6135(1987).			
RN	[3]			
RP	CHROMOSOMAL TRANSLOCATION.			
RX	MEDLINE; 94356859.			
RA	DUNN T., PRAISSMAN L., HAGAG N., VIOLA M.V.;			
RL	CANCER GENET. CITOGENET. 76:19-22(1994).			
RN	[4]			
RP	CHROMOSOMAL TRANSLOCATION.			
EX	MEDLINE; 94243799.			
RA	ICHIKAWA H., SHIMIZU K., HAYASHI Y., OHKI M.;			
RL	CANCER RES. 54:2865-2868(1994).			
CC	-1- SUBCELLULAR LOCATION: NUCLEAR.			
CC	-1- DISEASE: A FORM OF ACUTE MYELOID LEUKEMIA (AML) IS CHARACTERIZED			
CC	BY A CHROMOSOMAL TRANSLOCATION T(16;21)(P11;Q22) WHICH INVOLVES			
CC	ERG AND FUS.			
CC	-1- SIMILARITY: BELONGS TO THE ETS FAMILY.			
DR	EMBL; M17254; G182187; -.			
DR	EMBL; M21535; G182185; -.			
DR	PIR; A29515; TVHDEG. -.			
DR	TRANSFAC; T00265; -.			
DR	TRANSFAC; T00266; -.			
DR	MTM; 165080; -.			
DR	PROSITE; PS00345; ETS_DOMAIN_1; 1.			
DR	PROSITE; PS00346; ETS_DOMAIN_2; 1.			
DR	PROSITE; PS50061; ETS_DOMAIN_3; 1.			
KW	NUCLEAR PROTEIN; TRANSFORMING PROTEIN; DNA-BINDING;			
	ALTERNATIVE INITIATION; PROTO-ONCOGENE; CHROMOSOMAL TRANSLOCATION.			

FT	CHAIN	1	462	TRANSFORMING PROTEIN ERG-2.
FT	CHAIN	100	462	TRANSFORMING PROTEIN ERG-1.
FT	DNA_BIND	294	374	ETS-DOMAIN.
SO	SEQUENCE	462 AA:	52031 MW:	CECC2D2B CRC32:
Query Match 75.2%: Score 2608; DB 1; Length 462;				
Best Local Similarity 82.7%: Pred. NO. 0.00e+00;				
Matches 396; Conservative 27; Mismatches 29; Indels 27; Gaps 10.				
D	9	AAHKEALSVSEDSQSLFECAVGTPLHAKTEMTASSSSSDYQOTSKMSPRVPOQDMLSQPP	68	
O	2	ASTIEKALSVSEDSQSLFECAVGSPLHAKTEMTASSSSSEYQOTSKMSPRVPOQDMLSQPP	61	
D	69	ARVTKMKCNQVNGVSGNSPDECSYANGGKMGVNSPDYVGNNYSYMEKHMPPNNTTN	128	
O	62	ARVTKMKCNQVNGVSGNSPDECSYANGGKMGVNSPDYVGNNYSYMEKHMPPNNTTN	121	
D	129	ERRVYVPADPLMTSDHYHROWLENAVKEGADVINILFONIDKEICKMTKDDFORLTP	188	
O	122	ERRVYVPADPLMTSDHYHROWLENAVKEGADVINILFONIDKEICKMTKDDFORLTP	181	
D	189	SYNDAILLSHLHYLRETPPLHLTSDVDKALONSPRLMHAHNTD---L-P---Y-E--	236	
O	182	SYNDAILLSHLHYLRETPPLHLTSDVDKALONSPRLMHAHNTGATFIRPNISVYEAT	241	
D	237	-----P--P-----RSAMTGHGHPFOSKAAQSPSTVKTEDQRPQLDPRYQILGPTSS	284	
O	242	QRIITRPDLPEOARRSAMTSHSPT-OSKATOPSSSTVPKTEQRPQLDPRYQILGPTSS	300	
D	285	RLANGSGQIOLMOLFLELLSDSSNSGCIWEGNGEKKMTDPEVARRNGERSKPNMN	344	
O	301	RLANGSGQIOLMOLFLELLSDSSNSGCIWEGNGEKKMTDPEVARRNGERSKPNMN	360	
D	345	YDKLSRALRYYYDNKIMTKVHGKRYAYKDFHGIAQALQPHPESSLYKYP-SDLPYMG	403	
O	361	YDKLSRALRYYYDNKIMTKVHGKRYAYKDFHGIAQALQPHPESSLYKYP-SDLPYMG	419	
D	404	YHAPQKKNFAHPHPALPVTSSSFFAAPNPYNSPTGGIYVNTRLPTSHMPSHLGTY	462	
O	420	PHAPQKKNFAHPHPALPVTSSSFFAAPNPYNSPTGGIYVNTRLPTSHMPSHLGTY	478	
RESULT 2 STANDARD: PRT: 452 AA.				
ID	FL11_HUMAN	STANDARD:	PRT:	452 AA.
AC	001543;			
DT	01-JUL-1993 (REL. 26, CREATED)			
DT	01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)			
DT	01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)			
FT	FL1-1 ONCOGENE (ERB TRANSSCRIPTION FACTOR).			
FT	FL11.			
OS	HOMO SAPIENS (HUMAN).			
OC	EUMAROTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;			
OC	EUTHERIA; PRIMATES.			
NC	[1]			
NC	SEQUENCE FROM N.A.			
NC	TISSUE-BONE MARROW;			
NC	MEDLINE; 92396239.			
NC	DELAETRE O., ZUCMAN J., PLOUGASTEL B., DESMAZE C., MELOT T.,			
NC	PETER M., KOVAR H., JOUBERT I., DE JONG P., ROULEAU G.;			
NC	NATURE 359:162-165(1992).			
NC	[2]			
NC	SEQUENCE FROM N.A.			
NC	MEDLINE; 93079640.			
NC	WATSON D.K., SMYTH F.E., THOMPSON D.M., CHENG J.Q., TESTA J.R.,			
NC	PAPAS T.S., SETH A.;			
NC	CELL GROWTH DIFFER. 3:705-713(1992).			
NC	[3]			
NC	SEQUENCE FROM N.A.			
NC	MEDLINE; 93007976.			
NC	PRASAD D.D., RAO V.N., REDDY E.S.;			
NC	CANCER RES. 52:5833-5837(1992).			
NC	[4]			
NC	SEQUENCE FROM N.A.			

RC TISSUE=BLLOOD:
RA MEDLINE: 93176799.
RA HROMAS R., MAY W., DENNY C., RASKIND W., MOORE J., MAKI R.A.,
RA BECK E., KLEMSZ M.J.,
RL BIOCHIM. BIOPHYS. ACTA 1172:155-158(1993).
RN [5]
RP STRUCTURE BY NMR OF 276-373.
RX MEDLINE: 95292091.
RX LIANG H., MAO X., OLEJNICZAK E.T., NETTESHEIM D.G., YU L.,
RA MADGOWS R.P., THOMPSON C.B., FESIK S.W.,
RL NAT. STRUCT. BIOL. 1:871-875(1994).
CC -1- FUNCTION: SEQUENCE SPECIFIC TRANSCRIPTIONAL ACTIVATOR. RECOGNIZES
CC THE DNA SEQUENCE 5'C[CA]GAGACT-3'.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- DISEASE: A FORM OF EWING'S SARCOMA IS CHARACTERIZED BY A
CC CHROMOSOMAL TRANSLOCATION T(11;22)(Q24;Q12) WHICH INVOLVES FLI1
CC AND EWS.
CC -1- SIMILARITY: BELONGS TO THE ETS FAMILY.
DR EMBL: X67001: G32530.
DR EMBL: M98833: G182189.
DR EMBL: S45205: G257354.
DR EMBL: M93255: G182660.
DR PDB: 1FLI: 15-SEP-95.
DR TRANSFAC: T02066.
DR MIM: 193067.
DR PROSITE: PS00345; ETS_DOMAIN_1; 1.
DR PROSITE: PS00346; ETS_DOMAIN_2; 1.
DR PROSITE: PS50061; ETS_DOMAIN_3; 1.
KW TRANSCRIPTION REGULATION; ACTIVATOR; DNA-BINDING; NUCLEAR PROTEIN;
KW PROTO-ONCOGENE; CHROMOSOMAL TRANSLOCATION; 3D-STRUCTURE.
FT DNABIND 281 361 ETS_DOMAIN.
FT CONFLICT 69 69 E -> V (IN REF. 4).
FT CONFLICT 77 77 MISSING (IN REF. 4).
FT CONFLICT 130 130 P -> A (IN REF. 4).
FT CONFLICT 133 133 W -> V (IN REF. 4).
FT CONFLICT 323 323 E -> Q (IN REF. 3 AND 4).
FT CONFLICT 391 391 Y -> Q (IN REF. 2).
FT CONFLICT 426 426 MISSING (IN REF. 2 AND 4).
SQ SEQUENCE 452 AA; 50982 MW; 192476F7 CRC32;

Query Match 51.2% Score 1776; DB 1; Length 452:
Best Local Similarity 57.4%; Pred. No. 0.00e+00; Indels 38; Gaps 23:
Matches 278; Conservative 86; Mismatches 82;

Db 1 MDGTEKALSVSDQSILFSDYGAHAALPKRADTAGSPDYGQPHKINLPQOEWINQ 60
1 MASTIKELSVSDEQSLFECAVGP-HLAKTEMTASSSEYGTGSKMSRVYQQDMLSQ 59
Db 61 P-VRYNAREYD--HMNGRESPYDCSVSKCSKIVGGESNPNNYNSYDEKNGPPNN 117
P PPARVITIKMECPNQVNGSRNSPDDCSVAKGKGWSSSDVNGMYGSMYMEKH-IPPNM 118
Db 118 TTNERRYVVPADPLMTQOEHRQWLEAKIKEYSLMELDTSFQNMCKELCKNKKEFLR 177
TTNERRYVVPADPLMTQOEHRQWLEAKIKEYSLMELDTSFQNMCKELCKNKKEFLR 178
Db 178 ATTLETEVLLSHLYSRESL--LAYNTSHT-DOSSRL-SYE---D---P--S-Y- 222
179 LTPEYNADILLSHLYREPLPLRLTSDVDKALQNSPRMLHARNIGCATFIIPNTSYR 238
Db 223 DSV-R--R-G-ANGNNNGS-LNKS-PPLG-GAQ---TISKTEQRPQDPDYQILGPT 269
EATQRTTRPPLPEQARRSRKMTSHSHPTQSKATQPOSSSTVPRTEQRPQDLPDYQILGPT 298
Db 270 SSRLANGSGOIQMOFLLELLSSANAASCTTEGTGGEKMTDPDEVARMEERKSKPN 329
SSRLANGSGOIQMOFLLELLSSANSNCITWEGTGEKMTDPDEVARMEERKSKPN 358
Db 330 MNYDKLSRALYYYDKNIMTKVHCKRVAYKEDFGIAQALQAPHTSSMKYK-P-SDISYM 388
MNYDKLSRALYYYDKNIMTKVHCKRVAYKEDFGIAQALQAPHTSSMKYK-P-SDISYM 417
Db 389 PSYHAQKQVYVPPHDSMPVTSSEFFGAASQYWTSPGTGIPNPVNPVPHPTVHPSHL 448

Y	418	LPQNAHPQKNMFAVPHHPALPVLVTSSFFAANPYNWSPNTGTYINT---	RLPNAHMSHL	474
Db	449	GSYY 452		
Y	475	GTYY 478		
RESULT	3	STANDARD:	PRF:	452 AA.
ID	FL11.MOUSE			
AC	P26323;			
DT	01-MAY-1992 (REL. 22, CREATED)			
DT	01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)			
DT	01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)			
DE	RETROVIRAL INTEGRATION SITE PROTEIN FL1-1.			
GN	FL11 OR FL1-1.			
OS	MUS MUSCULUS (MOUSE).			
OC	EUDAROTA: METAZOA: CHORDATA: VERTEBRATA; TETRAPODA; MAMMALIA;			
OC	EUTHERIA: RODENTIA.			
RN	[1]			
RE	SEQUENCE FROM N.A.			
RA	STRAIN-BALB/C; TISSUE--SPLEEN;			
RA	MEDLINE; 91257578.			
RA	BEN-DAVID Y., GIDDENS E.B., LETWIN K., BERNSTEIN A.,			
RA	GENES DEV. 5:908-918(1991).			
RL	[2]			
RN	CHARACTERIZATION.			
RP	MEDLINE; 93275657.			
RX	ZHANG L., LEHARCHANDEL V., ROMEO P.-H., BEN-DAVID Y., GREER P.,			
RA	BERNSTEIN A.;			
RL	ONCOGENE 8:1621-1630(1993).			
CC	-1- FUNCTION: SEQUENCE SPECIFIC TRANSCRIPTIONAL ACTIVATOR. RECOGNIZES			
CC	THE DNA SEQUENCE 5'CA[CA]GGAGT-3'.			
CC	-1- SUBCELLULAR LOCATION: NUCLEAR.			
CC	-1- TISSUE SPECIFICITY: EXPRESSED IN BOTH HEMATOPOIETIC AND			
CC	NONHEMATOPOIETIC TISSUES.			
CC	-1- INDUCTION: SPECIFICALLY UP-REGULATED IN LEUKEMIC CLONES WITH			
CC	F-MUZY INSERTIONS UP-STREAM OF THE FL1-1 LOCUS.			
CC	-1- DISEASE: INVOLVED IN ERYTHROLEUKEMIA INDUCTION BY FRIEND MURINE			
CC	LEUKEMIA VIRUS (F-MUZY).			
CC	-1- SIMILARITY: BELONGS TO THE ETS FAMILY.			
CC	EMBL; X59421; G50975; -.			
DR	PIR; S17403; S17403.			
DR	TRANSFAC; T01408; -.			
DR	MED; MG1:95554; FL11.			
DR	PROSITE; PS00345; ETS_DOMAIN_1; 1.			
DR	PROSITE; PS00346; ETS_DOMAIN_2; 1.			
DR	PROSITE; PS50061; ETS_DOMAIN_3; 1.			
DR	TRANSCRIPTION REGULATION; ACTIVATOR; DNA-BINDING; NUCLEAR PROTEIN;			
DR	PROTO-ONCOGENE.			
SO	DNA_BIND	281	361	ETS-DOMAIN.
SE	SEQUENCE	452 AA;	51002 MW;	CECAGDD4 CRC32;
Query Match		51.1%;	Score 1773;	DB 1; Length 452;
Best Local Similarity		57.2%;	Pred. No. 0.00e+00;	
Matches	277;	Conservative	89;	Mismatches 80; Indels 38; Gaps 24;
Db	1	MDGTTKEALSVSDQSIFDSATYGAHAHLPRADMTAGSGPDYGPQPHKINLPQOEINQ	60	
Y	1	MASTIKELSVSEDSOLFECAYGSP-HIAKTEMTASSSSSEYCGTSSKSPVPODWLSQ	59	
Db	61	P-VARNVAREYD--HMNSRESPPDCSVSKCNKLVGGSEAPMNYNSYMEKNGPPPM	117	
Y	60	PPARVTIMECPNPNVNGSRNSPDDCSYAKGCKAVSSSDNAGMYGSMYMEKH-IPPPM	118	
Db	118	TTNERIVVPADPTLMTQEHVROWLEWAIKEYGLMEIDTSEFQNMDDKEJLCKRANKDEF	177	
Y	119	TTNERIVVPADPTLMTSDVHROWLEWAIKEYGLPVDILLFQNDKEJLCKRANKDEF	178	
Db	178	ATSVANVETVLSHLSYBESSL-LAYNTSHT-DQSSRL-NVKE-----D-----P-S-Y-	222	
Y	179	LPSTNADILISHLAHYLETPPLPHTSDVDKALQNSPRLNHAHNGATFIPNTSYVP	238	

Db	223	DSV-R--R-G-ANNNNNSG-LKRS-PLLG-GSG----	TMKNTEDRPPDPYQILGPT	269
Qy	229	EATGRITTRPDLPLEQARRSAMTSHSPHOTOSKATQBPSSSTVPRKTEDORPDLDPYQILGPT		298
Db	270	SSRLANPQSGQIQLMQFLELLLELLSDANASCIETWCTNGEETKMTDPDEVARMRGERSKPN		329
Qy	299	SSRLANPQSGQIQLMQFLELLLELLSDANASCIETWCTNGEETKMTDPDEVARMRGERSKPN		358
Db	330	MNYDKLSAALAYYYDKNIMTKYHCKRRAIKYFDFHGIQALQPHPTETSMKYP-SDSIYM		368
Qy	359	MNYDKLSAALAYYYDKNIMTKYHCKRRAIKYFDFHGIQALQPHPTETSMKYP-SDSIYM		417
Db	369	PSYHAAKQVNFVPSHSSSPMTSSSEFGAASQVMTSPTAGIYPNPSPHPTVHPSHL		448
Qy	418	LQPAHHPQKMFVAPHPALPVTSSSEFFAAPNPYNSPTGCTIPN-I--RLPAHHPSHL		474
Db	449	GSTY 452		
Qy	475	GTY 478		
RESULT	4	STANDARD;	PRT;	453 AA.
ID	AC	FL11_XENLA	P41157;	
DT	01-FEB-1995	(REL. 31, CREATED)		
DT	01-FEB-1995	(REL. 31, LAST SEQUENCE UPDATE)		
DT	01-FEB-1995	(REL. 31, LAST ANNOTATION UPDATE)		
DE	REPROVIRAL INTEGRATION SITE PROTEIN FL1-1 HOMOLOG.			
GN	FL1.			
OS	XENOPUS LAEVIS (AFRICAN CLAWED FROG).			
OC	EUROCYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AMPHIBIA; ANURA.			
RN	[1]			
RP	SEQUENCE FROM N. A.			
RX	MEULINE; 94206844.			
RA	MEYER D., WOLFE C. M., STEIGLER P., SENAN F., BEFORT N.,			
RA	BEFORT J. J., REMY P.,			
RL	MECH. DEV. 44:109-121(1993).			
CC	-1- SUBCELLULAR LOCATION: NUCLEAR.			
CC	-1- DEVELOPMENTAL STAGE: EXPRESSED DURING EMBRYOGENESIS.			
CC	-1- SIMILARITY: BELONGS TO THE ETS FAMILY.			
DR	EMBL; G56979; G505487; -.			
DR	TRANSFAC; T02067; -.			
DR	PROSITE; PS00345; ETS DOMAIN 1; 1.			
DR	PROSITE; PS00346; ETS DOMAIN 2; 1.			
DR	PROSITE; PS50061; ETS DOMAIN 3; 1.			
KW	TRANSCRIPTION REGULATION; ACTIVATOR; DNA-BINDING; NUCLEAR PROTEIN.			
FT	DNA_BIND 282 362 ETS-DOMAIN.			
SQ	SEQUENCE 453 AA; 51015 MW; DEB5A90B CRC32;			
Query Match	50.2%;	Score 1740;	DB 1;	Length 453;
Best Local Similarity	56.4%;	Pred. No. 0.00e+00;		
Matches	273;	Conservative	91;	Mismatches 83; Indels 37; Gaps 25.
Db	1	MDGTKEALSVSDQSIFDSYGASSHLKADMTAGANPGPHKINIPPOQDINQ		60
Qy	1	MASTIKELSVSDEQSLFECAVC-SPLHAKTEMTASSSEYGTGKSMSPRYQDMLSQ		59
Db	61	P-MRYNIRKREY--HMNGSRSPVDCSINCKSKILIGSEGNAMTY--TYMDEKNGPPPNM		116
Qy	60	PPARTITMECNPNQVNGSRNSPDDCSVAKGKMYSSDDVNGKSYGMEKH-IPPPNM		118
Db	117	TNTERVYVPADPLMSODHROKLEMAIKREYGEVLEIDSLFQINDEKELCKNKDEFLR		176
Qy	119	TNTERVYVPADPLMSODHROKLEMAIKREYGEVLEIDSLFQINDEKELCKNKDQDFOR		178
Db	177	TSIYNTIEVLSHTNLYLRDSS--SLIGYNTQANT--DQSSRLT-AKE---D---P--S-Y-		222
Qy	179	LTPSYNADILSHLYLRETPDLPLHLSDDVDKALQNSPRLMHARNNGATFIPPNISYVP		238
Db	223	EAV-R--RSG--WGNKSSP-VKSSPM-G-GTQ-NVAK-S-GDOORSPDPYQILGPT		270
Qy	239	EATGRITTRPDLPLEQARRSAMTSHSPHOTOSKATQBPSSSTVPRKTEDORPDLDPYQILGPT		298

ID	ERG_LYTVA	STANDARD:	PRT:	173 AA.
Db	271	SSANPSSGGIOIOMQFLELLSSSSNASCITMGTEGFEKMTDPDVARRMGRKSKPN	330	
Qy	299	SSRLANPSSGGIOIOMQFLELLSSSSNASCITMGTEGFEKMTDPDVARRMGRKSKPN	358	
Db	331	MNYDKLSRLALYYVYDKSIMTKVGRKRYAKKDFEGIAOALOPHTDPSMYKP-SSESYM	389	
Qy	359	MNYDKLSRLALYYVYDKSIMTKVGRKRYAKKDFEGIAOALOPHTDPSMYKP-SSESYM	417	
Db	390	PSYSHOOKVNFVPSHSPVYSSGFEFGATSPYWNSSPANITYPNVPRHPNTHVOSHL	449	
Qy	418	LQPAHAPKQKMFVAPHPALPVTSSSFEFAAPNPWNSPTGGIYENT--RLPAHMPSHL	474	
Db	450	GGFY 453		
Qy	475	GTYY 478		
RESULT	5	STANDARD:	PRT:	272 AA.
ID	ERG_MOUSE			
Db	15-JUL-1998 (REL. 36, CREATED)			
Qy	15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)			
Db	15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)			
Qy	15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)			
Db	TRANSFORMING PROTEIN ERG (FRAGMENT).			
Qy	ERG OR ERG-3			
Db	MUS MUSCULUS (MOUSE).			
Qy	EURAROTA, METAOCIA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;			
Db	EUTHERIA; RODENTIA.			
Qy	[1]			
Db	SEQUENCE FROM N.A.			
Qy	MEDLINE; 94019387.			
Db	RIVERA R. R., STUTER M. H., STEENBERGEN R., MURRE C.;			
Qy	MOL. CELL. BIOL. 13:7163-7169(1993).			
Db	-1- SUBCELLULAR LOCATION: NUCLEAR.			
Qy	-1- SIMILARITY: BELONGS TO THE ETS FAMILY.			
Db	EMBL: S66169; GA36277; -.			
Qy	MED. MGT:95415; ERG.			
Db	PROSITE; PS00345; ETS_DOMAIN_1; 1.			
Qy	PROSITE; PS00346; ETS_DOMAIN_2; 1.			
Db	PROSITE; PS50061; ETS_DOMAIN_3; 1.			
Qy	NUCLEAR PROTEIN; TRANSFORMING PROTEIN; DNA-BINDING; PROTO-ONCOGENE.			
Db	NON_TER 1			
Qy	FT DNA_BIND 138 218 ETS-DOMAIN.			
Db	FT NON_TER 272			
Qy	SEQUENCE 272 AA; 30907 MW; D2C258D9 CRC32;			
Query March	45.3%	Score 1571; DB 1; Length 272;		
Best Local Similarity	82.4%	Pred. No.0.0e+00;		
Matches 22;	Conservative 17; Mismatches 28; Indels 3; Gaps 3;			
Db	1	DDFOGLTSPYVADILSLHLHYLRETPHLRLSDVDKALQNSPLHMAHARTGGAFTFPN	60	
Qy	174	DDFOGLTSPYVADILSLHLHYLRETPHLRLSDVDKALQNSPLHMAHARTGGAFTFPN	233	
Db	61	TSVPEATORTITRPDLPEPPRRSASMTGSHLTPQSKAQPSPASVAKTEDQRPQDLPY	120	
Qy	234	TSVPEATORTITRPDLPEPPRRSASMTGSHLTPQSKAQPSPASVAKTEDQRPQDLPY	292	
Db	121	QILGTSRLANPSSGGIOIOMQFLELLSSSSNASCITMGTEGFEKMTDPDVARRMGR	180	
Qy	293	QILGTSRLANPSSGGIOIOMQFLELLSSSSNASCITMGTEGFEKMTDPDVARRMGR	352	
Db	181	RKSPNNMYDKLSRLALRYVYDKSIMTKVGRKRYAKKDFEGIAOALOPHTDPSMYKP	239	
Qy	353	RKSPNNMYDKLSRLALRYVYDKSIMTKVGRKRYAKKDFEGIAOALOPHTDPSMYKP	411	
Db	240	SDLPYMGSYHAHPQKMFVSPHPALPVTSSSF 272		
Qy	412	HGIAOALOPHAHPQKMFVADHPALPVTSSSF 444		
RESULT	6	STANDARD:	PRT:	173 AA.
ID	ERG_LYTVA			

Query Match	Best Local Similarity	Score	DB 1	Length	DB 2	Score	DB 3	Length
Query Match	Best Local Similarity	Score	DB 1	Length	DB 2	Score	DB 3	Length
77.6%	Pred. No. 1,25e-89							
15.0%	Score 519; DB 1; Length 102;							

[illegible]

```

RX MEDLINE 90356411.
RA MOFF C.M., STIEGLER P., BALTZINGER M., MEYER D., GHYSDAEL J.,
RA STEHELIN D., BEFORT N., REMY P.;
RL NUCLEIC ACIDS RES. 18:4603-4604(1990).
RN [3]
RP SEQUENCE OF 121-472 FROM N.A.
RX MEDLINE 92088972.
RA MOFF C.M., STIEGLER P., BALTZINGER M., MEYER D., GHYSDAEL J.,
RA STEHELIN D., BEFORT N., REMY P.;
RL CELL GROWTH DIFFER. 2:447-456(1991).
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- SIMILARITY: BELONGS TO THE ETS FAMILY.
DR EMBL: M81683: G214130: -.
DR EMBL: X51826: G64619: -.
DR PIR: S10994: S10994.
DR PIR: S28824: S28824.
DR TRANSFAC: T02041: -.
DR PROSITE: PS00345: ETS_DOMAIN_1; 1.
DR PROSITE: PS00346: ETS_DOMAIN_2; 1.
DR PROSITE: PS50061: ETS_DOMAIN_3; 1.
RW DNA-BINDING; NUCLEAR PROTEIN.
FT DNA_BIND 366 446 ETS-DOMAIN.
SO SEQUENCE 472 AA; 53894 MW; 064d44ea crc32;

Query Match 13.1%; Score 453; DB 1; Length 472;
Best Local Similarity 59.2%; Pred. No. 3,63e-74;
Matches 61; Conservative 16; Mismatches 24; Indels 2; Gaps 2;

Db 354 ASLAFGTSGPIDLMQFLLELLDKSCGSFISWTGDMGEFKLDPDEVARMRGKRANKP 413
OY 299 SSRLAN-PGSGQIDLMQFLLELLSDSSNSCITWEGTNGEFGKFLDPDEVARMRGKRANKP 357
Db 414 KMYEKLRLGLRYDDNIINIIHKT-SGKRYVRYFVCDLHNLGY 455
OY 358 NMNYDKSLRALRYDDKIMTKVHPPESSMYKKPSPDLPLYMSY 400

RESULT 15
ID ETS_A_CHICK STANDARD; PRT; 441 AA.
AC P13474;
DT 01-JAN-1990 (REL. 13, CREATED)
DT 01-JAN-1990 (REL. 13, LAST SEQUENCE UPDATE)
DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
DE TRANSFORMING PROTEIN P54/C-ETS-1.
GN ETS-1.
OS GALLUS GALLUS (CHICKEN).
OC ENKARYOTA; METAZOA; CHORATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
OC GALLIFORMES.
CC [1]
RN SEQUENCE FROM N.A.
RP TISSUE-SPLEEN;
RX MEDLINE 88289026.
RA DUTERQUE-COQUILLIARD M., LEPRINCE D., FLOURENS A., HENRY C.,
RA GHYSDAEL J., DEBUIRE B., STEHELIN D.;
RL ONCOGENE RES. 2:335-344(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-SPLEEN;
RX MEDLINE 88289029.
RA CHEN J.H.;
RL ONCOGENE RES. 2:371-384(1988).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE 89045631.
RA WATSON D.K., MCHILLIAMS M.J., PAPAS T.S.;
RL VIROLOGY 167:1-7(1988).
CC -1- FUNCTION: THIS PROTEIN IS THE NORMAL CELLULAR PRODUCT OF CHICKEN
CC ETS. IN THE E26 VIRUS, ETS IS RESPONSIBLE FOR ERYTHROBLAST
CC TRANSFORMATION.
CC -1- TISSUE SPECIFICITY: THYMUS.
CC -1- ALTERNATIVE PRODUCTS: ALTERNATIVE SPLICING MECHANISM WITHIN THE
CC C-ETS-1 GENE GENERATES TWO PROTEINS HAVING DISTINCT N-TERMINAL
CC EXTREMITY: P54 AND P68.

```

NWSEKTH (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Search_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Apr 7 09:32:06 1999; MasPar time 19.45 Seconds

Abular output not generated.

Title: >US-08-878-177-2
Description: (1-451) from US08878177.pep
Perfect Score: 3286
Sequence: 1 MASTIKELSVSEDSLEF.....IYPTRLPAHMPHILGTRY 451

Scoring table: PAM 150
Gap 11

Searched: 116695 segs, 37453910 residues

Post-processing: Minimum Match 08
Listing first 45 summaries

Database: p1r58
1: p1r1 2: p1r2 3: p1r3 4: p1r4

Statistics: Mean 48.029; Variance 92.580; scale 0.519

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	2989	91.0	478	1	S60754	transcription factor 0.00e+00
2	2887	87.9	486	1	TVHUEG	transcription factor 0.00e+00
3	2348	71.5	452	1	I37565	transcription factor 0.00e+00
4	2339	71.2	452	1	S17403	transcription factor 0.00e+00
5	2305	70.1	453	1	S49013	transcription factor 0.00e+00
6	1973	60.0	385	1	S29844	transcription factor 0.00e+00
7	1535	46.7	272	2	A54617	transcription factor 8.79e-295
8	959	29.2	173	2	A56646	transcription factor 6.93e-171
9	895	27.2	179	2	S51226	transcription factor 3.51e-157
10	660	20.1	464	1	S37616	transcription factor 6.37e-107
11	653	19.9	454	1	A40858	GA-binding protein al 4.38e-105
12	649	19.8	454	1	A48146	nuclear respiratory f 1.19e-99
13	640	19.5	454	2	I38739	transcription factor 8.80e-73
14	623	19.0	103	2	S28823	transcription factor 1.41e-72
15	493	15.0	211	2	S24300	transcription factor 3.61e-72
16	492	15.0	268	2	S11224	transcription factor 3.61e-72
17	492	15.0	438	1	S11225	transcription factor 3.61e-72
18	490	14.9	440	1	A35875	transcription factor 3.61e-72
19	490	14.9	441	1	A53988	transcription factor 3.61e-72
20	490	14.9	441	1	TVHUEE	transcription factor 3.61e-72
21	486	14.8	441	1	TVCHTE	transcription factor 2.36e-71
22	487	14.8	472	1	B53236	transcription factor 1.48e-71
23	486	14.8	485	1	TVCHET	transcription factor 2.36e-71

24	485	14.8	623	1	S33167	pointed protein, spli	3.78e-71
25	485	14.8	718	1	S33168	pointed protein, spli	3.78e-71
26	483	14.7	468	1	TVHSE2	transcription factor	9.67e-71
27	483	14.7	469	1	TVHUE2	transcription factor	9.67e-71
28	484	14.7	472	1	A53236	transcription factor	6.05e-71
29	482	14.7	478	1	TVHSE2	transcription factor	1.53e-70
30	480	14.6	479	1	TVCHHE2	transcription factor	3.95e-70
31	474	14.4	110	2	A45938	transcription factor	6.59e-69
32	469	14.3	548	2	S59133	ETS2 repressor factor	6.85e-68
33	467	14.2	94	2	S28820	transcription factor	1.75e-67
34	462	14.1	440	2	I48291	transcription factor	1.81e-66
35	460	14.0	477	1	I38893	transcription factor	4.60e-66
36	460	14.0	477	1	B46396	transcription factor	4.60e-66
37	456	13.9	250	2	A54308	PE-1 protein - human	2.98e-65
38	454	13.8	510	1	S43692	transcription factor	7.57e-65
39	443	13.5	342	2	A46396	ets-related protein 7	1.27e-62
40	443	13.5	462	1	S35534	adenovirus E1A enhanc	1.27e-62
41	443	13.5	555	1	S24061	transcription factor	1.27e-62
42	437	13.3	428	1	TVHUEK	transcription factor	2.06e-61
43	431	13.1	429	2	JC4965	elk1 protein - mouse	3.34e-60
44	429	13.1	430	2	I48755	MSAP1 - mouse	8.44e-60
45	422	12.8	407	2	I38062	Net - human	2.15e-58

ALIGNMENTS

RESULT	ENTRY	1	ALIGNMENTS
	S60754	#type complete	
	transcription factor erg - chicken		
	#formal_name Gallus gallus #common_name chicken		
	DATE	27-Apr-1996 #sequence__revision 23-May-1997 #text__change 10-Jul-1998	

ACCESSIONS	REFERENCE
S60754	Dhordain, P.; Dewitte, F.; Desblens, X.; Stehelin, D.;
S60754	Dutergue-Cogulland, M.
	Mech. Dev. (1995) 50:17-28
	Mesodermal expression of the chicken erg gene associated with
	precartilaginous condensation and cartilage
	differentiation.

#cross-references	MOID:95329425
#accession	S60754
#status	preliminary
#molecule_type	mRNA
#residues	1-478 #label DHO
GENETICS	#cross-references EMBL:X77159; NID:g790439; PID:g790440

CLASSIFICATION	erg
#gene	superfamily transcription factor erg; ets DNA-binding domain
KEYWORDS	homology; ets RII regulatory region homology
	DNA binding; phosphoprotein; proto-oncogene; transcription
	factor; transforming protein

FEATURE

119-193	#domain ets RII regulatory region homology #label ETS2\
312-390	#domain ets DNA-binding domain homology #label ETS

SUMMARY

#length 478	#molecular-weight 53913	#checksum 547
-------------	-------------------------	---------------

Query Match

Best Local Similarity	91.08; Score 2989; DB 1; Length 478;
Matches	451; Conservative 0; Mismatches 0; Indels 27; Gaps 2;

DB	1	MASTIKELSVSEDSLEFCAGSPHLAKTEMTASSSSFGTSMSPVPODMLISOP	60
QY	1	MASTIKELSVSEDSLEFCAGSPHLAKTEMTASSSSFGTSMSPVPODMLISOP	60

DB

61	PARVITIMECNPQVNGSRNSPDDCSYAKGKMYSSSDNNGMYGMEKH1PPNMTT	120	
QY	61	PARVITIMECNPQVNGSRNSPDDCSYAKGKMYSSSDNNGMYGMEKH1PPNMTT	120

DB

121	NERRVIVPADPTLMSIDHVRQMLEMAVKEYGLPDVDILLFQNDGKELCKMTKDDFORLT	180	
QY	121	NERRVIVPADPTLMSIDHVRQMLEMAVKEYGLPDVDILLFQNDGKELCKMTKDDFORLT	180

Db	181	PSYNDILLSHLHTLREPLPHLTSDDVDKALQNSPLRMHNRNGATFTFPNTSVYPEA	24.0
Qy	181	PSYNDILLSHLHLRE-----+-----R-----GATTFPNTSVYPEA	21.3
Db	241	TQRTITRPDLPEQARRSAMTSHSHPTQSKATOPSSSTVPEKTEDQRPOLDPYQILGPTSS	30.00
Qy	214	TQRTITRPDLPEQARRSAMTSHSHPTQSKATOPSSSTVPEKTEDQRPOLDPYQILGPTSS	27.3
Db	301	RLANGSQIOLMOLFLELLSDSSNSNCITTEGNGEFGKMTDPDEVARRNGERKSFRNMN	36.66
Qy	274	RLANGSQIOLMOLFLELLSDSSNSNCITTEGNGEFGKMTDPDEVARRNGERKSFRNMN	33.33
Db	361	YDKLSRALRYYYDKNIMTKVHGKRAYAKFEDEHGIQAOLQHPPESSMYKYPSDLPLYMSY	42.0
Qy	334	YDKLSRALRYYYDKNIMTKVHGKRAYAKFEDEHGIQAOLQHPPESSMYKYPSDLPLYMSY	39.33
Db	421	HAHPQKNMFVAPHPALPVTSSSFEAAPNPWNSPTGTYNTEBLPAAHMPSHLGTTY	47.8
Qy	394	HAHPQKNMFVAPHPALPVTSSSFEAAPNPWNSPTGTYNTEBLPAAHMPSHLGTTY	45.1

```

RESULT 2
ENTRY TITLE TVHUEG #type complete
CONTAINS transforming protein erg-3 - human
ORGANISM transforming protein erg-1: transforming protein erg-2
DATE #format_name Homo sapiens #common_name man
31-Mar-1989 #sequence_revision 30-May-1997 #text_change
31-Oct-1997

ACCESSIONS A94294; A94178; I58410; A28041; A29515
REFERENCE
#authors Rao, V.N.; Papas, T.S.; Shyam, E.; Reddy, P.
#journal Science (1987) 237:635-639
#title erg, a human ets-related gene on chromosome 21: alternative
splicing, polyadenylation, and translation.
#cross-references MUID:87263429
#accession A94294
##molecule_type mRNA
##residues 1-231,256-486 #label REA
#cross-references GB:M17254; NID:g182186; PID:g182187
REFERENCE A94178
#authors Reddy, E.S.P.; Rao, V.N.; Papas, T.S.
#journal Proc. Natl. Acad. Sci. U.S.A. (1987) 84:6131-6135
#title The erg gene is a human gene related to the ets oncogene.
#cross-references MUID:87317608
#accession A94178
##molecule_type mRNA
##residues 100-231,256-486 #label REB
#cross-references GB:M21535; NID:g182182; PID:g182185
REFERENCE I58410
#authors Prasad, D.D.; Rao, V.N.; Lee, L.; Reddy, E.S.
#journal Oncogene (1994) 9:669-673
#title Differentially spliced erg-3 product functions as a
transcriptional activator.
#cross-references MUID:94119611
#accession I58410
#status Preliminary: translated from GB/EMBL/DBJ
#molecule_type mRNA
##residues 230-259 #label RES
#cross-references GB:S68130; NID:g544960; PID:g544961
GENETICS
#gene GDB:ERG
#cross-references GDB:119884; OMIM:165080
#map_position 21q22.2-21q22.2
CLASSIFICATION #superfamily transcription factor erg; ets DNA-binding domain;
homology; ets RII regulatory region homology
KEYWORDS alternative splicing; DNA binding; nucleus; proto-oncogene;
transcription factor; transforming protein
FEATURE
1-231,256-486 #product transforming protein erg-2 #status predicted
100-231,256-486 #label EG2\
#product transforming protein erg-1 #status predicted
126-200 #label EG1\
#domain ets RII regulatory region homology #label ERS2\

```

320-998	#domain	ets	DNA-binding	domain	homology	#label	ETS
SUMMARY	#length	486	#molecular-weight	54608	#checksum	9577	
Query Match		87.9%	Score	2887;	DB 1;	Length	486;
Best Local Similarity		89.3%	Pred. No.	0.00e+00;			
Matches	427;	Conservative	20;	Mismatches	3;	Indels	28;
						Gaps	3

Db	9	AAKKEALSVSDOSLFECAVGTPLATETMTASSSSVGGTSGKMSRVRPOOMLSOP	68
Qy	2	ASTIKELALSVSDOSLFECAVGTPLATETMTASSSSVGGTSGKMSRVRPOOMLSOP	61
Db	69	ARVATKMECPNOCVNSRNSPDECSYAKGKMYGSPDTVGNMYGSYMEKKHMPNNMTN	128
Qy	62	ARVATKMECPNOCVNSRNSPDECSYAKGKMYGSPDTVGNMYGSYMEKKHMPNNMTN	121
Db	129	ERRVYVADPTLSTSTHVQMLEMANKEGEDVWILLFQNTDCKELCKMKKDDFORLTP	188
Qy	122	ERRVYVADPTLSTSTHVQMLEMANKEGEDVWILLFQNTDCKELCKMKKDDFORLTP	181
Db	189	SYNADILSLHLRLTPTPHLITSDVDVLCALONSPRLMHRNMGAAEIPFNTSVPEAT	248
Qy	182	SYNADILSLHLRLTPTPHLITSDVDVLCALONSPRLMHRNMGAAEIPFNTSVPEAT	214
Db	249	QRITTRPDLPEPPRRSANTGCHGPPPOSKAAOPSPSTVPKTEDORPOLDYQLGPTSS	308
Qy	215	QRITTRPDLPEPPRRSANTGCHGPPPOSKAAOPSPSTVPKTEDORPOLDYQLGPTSS	273
Db	309	RLANPGSGQQLQMLQFLLELLSDSNSCITTWGCTGEEFKMTDPDEVARMEERKSKPMN	368
Qy	274	RLANPGSGQQLQMLQFLLELLSDSNSCITTWGCTGEEFKMTDPDEVARMEERKSKPMN	333
Db	369	YDKLSRLARYYDKNIMTKVHGKRAYKEDFHGIALQALOPHPRESSLYKPYSDLPYGSY	428
Qy	334	YDKLSRLARYYDKNIMTKVHGKRAYKEDFHGIALQALOPHPRESSLYKPYSDLPYGSY	393
Db	429	HAHQKKNFAAPRPPALPTSSSFFAARPYNNSPTGGIYPTRTLPTSHMSHLGTY	486
Qy	394	HAHQKKNFAAPRPPALPTSSSFFAARPYNNSPTGGIYPTRTLPTSHMSHLGTY	451
RESULT	3	137565 #type complete	
ENTRY		transforming protein fl1, long splice form - human	
TITLE		Friend leukemia integration protein 1; transcription factor	
ALTERNATE_NAMES		ERGB	
ORGANISM		#formal_name Homo sapiens #common_name man	
DATE		04-Oct-1996 #sequence_revision 30-May-1997 #text_change	
ACCESSIONS		137565: S29843: S35506: A49000: A49015: I54170	
REFERENCE		137378	
#authors		Delattre, P.; Zucman, J.; Piongas, B.; Desmaziere, C.; Molot,	
		T.; Peter, M.; Kovar, H.; Jobert, I.; de Jong, P.;	
		Rouleau, G.; Aurias, A.; Thomas, G.	
#journal		Nature (1992) 359:163-165	
#title		Gene fusion with an ETS DNA-binding domain caused by	
		chromosome translocation in human tumours.	
#cross-references		WID:92396239	
#accession		I37565	
#status		preliminary: translated from GB/EMBL/DBJ	
##molecule_type		mRNA	
##residues		1-452 ##label RES	
##cross-references		EMBL:X67001; NID:g32529; PID:g32530	
REFERENCE		S29843	
#authors		Homas, R.; May, W.; Denny, C.; Roskind, W.; Moore, J.; Makl,	
		R.A.; Beck, E.; Klemz, M.J.	
#journal		Biochim. Biophys. Acta (1993) 1172:155-158	
#title		Human FLI-1 localizes to chromosome 11Q24 and has an aberrant	
		transcript in neuroepithelioma.	
#cross-references		WID:93176799	
#accession		S29843	
##molecule_type		mRNA	
##residues		1-66, 'V', '70-129', 'A', '131-132', 'V', '134-322', 'O', '324-425',	
#residues		427-452 ##label HNOI	


```

#cross-references EMBL:M93255
#note
    The authors translated the codon GTG for residue 69 as
    glu

REFERENCE
#authors
    Hromas, R.A.; May, N.; Denny, C.; Reskind, W.; Moore, J.;
    Maki, R.A.; Beck, E.; Klemz, M.J.
#submissions
    submitted to the EMBL Data Library, May 1993
#description
    Human Fli-1, an ETS oncogene family member preferentially
    expressed in hematopoietic cells, is a sequence-specific DNA
    binding protein.

#accession
    S35506
#molecule_type
    mRNA
#residues
    1-88, 'V', 70-76, 78-129, 'A', 131-132, 'V', 134-322, 'Q',
    324-425, 427-452 ##label RH02
#cross-references EMBL:M93255; NID:g182659; PID:g182660
A49000
#authors
    Prasad, D.D.; Rao, V.N.; Reddy, E.S.
#journal
    Cancer Res. (1992) 52:5833-5837
#title
    Structure and expression of human Fli-1 gene.
#cross-references MUID:93007976
#accession
    A49000
#status
    Preliminary
#molecule_type
    mRNA
#residues
    1-322, 'Q', 324-452 ##label PRA
#cross-references GB:M8833; NID:g257353; PID:g257354
#note
    Sequence extracted from NCBI backbone (NCBIN:115336,
    NCBI:P:115337)

REFERENCE
#authors
    Watson, D.K.; Smyth, F.E.; Thompson, D.M.; Cheng, J.Q.;
    Testa, J.R.; Papas, T.S.; Seth, A.
#journal
    Cell Growth Differ. (1992) 3:705-713
#title
    The ERGB/Fli-1 gene: isolation and characterization of a new
    member of the family of human ETS transcription factors.
#cross-references MUID:93075640
#accession
    A49015
#status
    Preliminary
#molecule_type
    mRNA
#residues
    1-390, 'Q', 392-425, 427-452 ##label WAT
#cross-references GB:M8833; NID:g182188; PID:g182189
#note
    Sequence extracted from NCBI backbone (NCBIN:119390,
    NCBI:P:119392)

REFERENCE
#authors
    Baud, V.; Lipinski, M.; Rassart, E.; Polluin, L.; Bergeron,
    D.
#journal
    Genomics (1991) 11:223-224
#title
    The human homolog of the mouse common viral integration
    region, Fli1, maps to 11q23-q24.
#cross-references MUID:92112219
#accession
    I54170
#status
    Preliminary; nucleic acid sequence not shown; translated
    from GB/EMBL/DBJ
#molecule_type
    DNA
#residues
    1-6 ##label RZ2
#cross-references GB:L47616; NID:g1000863; PID:g1000864

GENETICS
#gene
    GDB:Fli1
#cross-references GDB:127565; OMIM:193067
#map_position
    11q24.1-11q24.3
#introns
    77/3
CLASSIFICATION
#superfamily
    transcription factor erg; ets DNA-binding domain
#homology
    ets RII regulatory region homology
#alternative splicing
    DNA binding; nucleus; proto-oncogene;
    transcription factor; transforming protein

FEATURE
118-192
283-361
SUMMARY
#domain
    ets RII regulatory region homology #label ETS2\
    #domain
    ets DNA-binding domain homology #label ETS
#length
    452 #molecular-weight
    50982 #checksum
    389

Query Match
Best local similarity
Matches
    306; conservative
    85; mismatches
    55; indels
    11; gaps
    8;

1 MDGTRKALSVSDOSLFDLSAYGAANAHLPKADMTASGSPDYQCPHKINFLPPOEQWING 60

```

QY	1	MASTIKREALSVSEOSLSECAIYGSF-HIAKTEMTASSSEYQTSISMSPRVPOQDWLSQ	59
Db	61	P-VRNVKREEDY--HNGSRESVPDSCVSKSCSLVAGGSEPNPNYNSYMDERKNGPPNM	117
QY	60	PARPTIKHECPNNOVNGSRNSPDDSCVAKGGAMSSSDNVGNNGSYMEKH-IPPNM	118
Db	118	TNERRVIVPADPTIMTQEHVQWLEMAIKESLMEIDTSFQONMDKELCKNKKEDFLR	177
QY	119	TNERRVIVPADPTIMTQEHVQWLEMAIKESLMEIDTSFQONMDKELCKNKKEDFLR	178
Db	178	ATTLYNTEVLHSLSLRE-SSULAY-NTTSHTDQSSRLSVKXKDPSTYDVRGANGNNM	235
QY	179	LTPSYNADILHSHLYRERGAFTFIPNNTSYPEAQRITTRDEPLYEQARRSAMTSSH	238
Db	236	SGLNSPPLGAGQTSKNTKEORPQPPQYIIGTSSRLNPGSGQIQLOMFLLELSDSA	295
QY	239	PQSKATPOSSS-TYPKTDQRPODLPYQILGTSRLNPGSGQIQLOMFLLELSDSS	297
Db	296	NASCITWETNGEPMKTDDEVARBMGEKSKXPNMNYDKLSRLARYYYDKNIMTKVHGR	355
QY	298	NSNCITWETNGEPMKTDDEVARBMGEKSKXPNMNYDKLSRLARYYYDKNIMTKVHGR	357
Db	356	YAKKDFHGIQAOLQPHPESSMYKYPDSISYPSYHAQOQYNEVPHPSSMPTSSSF	415
QY	358	YAKKDFHGIQAOLQPHPESSMYKYPDSISYPSYHAQOQYNEVPHPSSMPTSSSF	417
Db	416	FGAASQYWTSPGTGIPNPNVPHRHTVPSHSGSY	452
QY	418	FAAPNPYNSPTGIGIYPTNT--RLPAHMPSHGITY	451
RESULT	4	S17403 #type complete	
ENTRY		transforming protein fl1 - mouse	
TITLE		Friend leukemia integration protein 1; transcription factor	
ALTERNATE_NAMES		ERB	
ORGANISM		Formal_name Mus musculus #common_name house mouse	
DATE		1-Dec-1993 #sequence_rev1 31-Dec-1993 #text_change	
ACCESSIONS		517403	
REFERENCE		S17403	
#authors		Ben-David, Y.; Giddens, E.B.; Letwin, K.; Bernstein, A.	
#journal		Genes Dev. (1991) 5:908-918	
#title		Erythroleukemia induction by Friend murine leukemia virus: insertional activation of a new member of the ets gene family, fl1-1, closely linked to c-ets-1.	
#cross-references		UID:91257578	
#accession		S17403	
#molecule_type		mrna	
#residues		1-452	
#cross-references		GB:X59421; NID:950974; PID:950975	
GENETICS			
#gene		fl1	
CLASSIFICATION		#superfamily transcription factor erg; ets DNA-binding domain	
KEYWORDS		homology: ets RII regulatory region homology DNA binding; nucleus; proto-oncogene; transcription factor; transforming protein	
FEATURE			
118-192		#domain ets RII regulatory region homology #label ETS2\	
283-361		#domain ets DNA-binding domain homology #label ETS	
SUMMARY		#length 452 #molecular-weight 51002 #checksum 8905	
Query Match		71.2%; Score 2339; DB 1; Length 452;	
Best Local Similarity		66.7%; Pred. No. 0.00e+00;	
Matches		305; Conservative 85; Mismatches 56; Indels 11; Gaps 9;	
Db	1	MDGITREALSVSDOSLSDSAVGAAGH.PKADMTASGSPDYGQPHKINLPPOEKGING	60
QY	1	MASTIKREALSVSEOSLSECAIYGSF-HIAKTEMTASSSEYQTSISMSPRVPOQDWLSQ	59
Db	61	P-VRNVKREEDY--HNGSRESVPDSCVSKCCKLVAGGSEPNPNYNSYMDERKNGPPNM	117

QY 60 PPARVTIKMECNNOYNGSNPDDCSVAKGKMWSSSDNVGNNGSYMEKH-IPPPM 118
Db 118 TNERRVIVPADPTLMTQEHVROWLEMAIKEYGLMEIDTISFFQNMKGELCKNKEDFLR 177
QY 119 TNERRVIVPADPTLMTSTHROWLEMAIKEYGLPDVDILLFQNDGKELCKMTKDDFOR 178
Db 178 ATSNVTEVLISHLSLRE-SSLLAY-NTSHTDQSSRLVAKEDPSYDVRGAMNNMN 235
QY 179 LTPSYNADILLSHLYLRGATFIEPNTSVPEATQRTTRDPLPEQARRSAMTSSH 238
Db 236 SGLNKPPLGSGOTMGKNTRQORPOPPYQILGPTSSRLANPSSGQIOLMQLLELSDA 295
QY 239 PTOAKATQSSS-TVPEKTEQORPOLDPYQILGPTSSRLANPSSGQIOLMQLLELSDS 297
Db 296 NASCITWEGTNGEFKMTDPDEVARRMGERKSKPMNNDKLSRLRYXXDNIMTKVHGR 355
QY 298 NSNCITWEGTNGEFKMTDPDEVARRMGERKSKPMNNDKLSRLRYXXDNIMTKVHGR 357
Db 356 YAKKDFPHGIAQALOPHPTISMRYKYPSEFSYMPYSHQOKVNPSPHSSMPTSSG 415
QY 358 YAKKDFPHGIAQALOPHPTISMRYKYPSEFSYMPYSHQOKVNPSPHSSMPTSSG 417
Db 416 FGAASQYMTSPTAGIYPNPSPHPTVPSHLSGYX 452
QY 418 FAAPNPYMNSTPGIYPTN-T-RLPAHMPSHLGTY 451
RESULT 5
ENTRY 549013 #type complete
TITLE transforming protein fl1 - African clawed frog
ALTERNATE_NAMES ERGB Friend leukemia integration protein 1; transcription factor
ORGANISM #formal_name Xenopus laevis #common_name African clawed frog
DATE 07-May-1995 #sequence_revision 23-May-1997 #text_change 05-Sep-1997
ACCESSIONS S49013
REFERENCE S49013
#authors Meyer, D.; Wolff, C.M.; Stiegler, P.; Senan, F.; Befort, N.;
#journal Befort, J.J.; Remy, P.
#title Mech. Dev. (1993) 44:109-121
#cross_references X1-fil, the xenopus homologue of the fl1-1 gene, is expressed during embryogenesis in a restricted pattern evocative of neural crest cell distribution.
#accession S49013
#status preliminary
#molecule_type mRNA
#residues 1-453 #label MEY
#cross_references EMBL:X66979; NID:g505486; PID:g505487
NETICS
#gene fl1
CLASSIFICATION #superfamily transcription factor erg; ets DNA-binding domain
KEYWORDS homology; ets RII regulatory region homology
DNA binding; nucleus; proto-oncogene; transcription factor;
transforming protein
FEATURE
117-191 #domain ets RII regulatory region homology #label ETS2\N
284-352 #domain ets DNA-binding domain homology #label ETS
SUMMARY #length 453 #molecular_weight 51015 #checksum 774
Query Match 70.1% Score 2305; DB 1; Length 453;
Best Local Similarity 65.9%; Pred. No. 0.00e+00;
Matches 302; Conservative 85; Mismatches 59; Indels 12; Gaps 9;
Db 1 MDGTEKALSVSDQSLFDSAYGASSHLSKADMTASANDYGOPHKINIPPOQDWINQ 60
QY 1 MASTIKELASVSEDQSLFECAYG-SPLAKTEMTASSSEYQGTSMSPRVQDMLSQ 59
Db 61 P-WRVVIRKREY--HMGSESPYDCKSINKSKLIGSEGNAMTY-TYDEKNGPPPPM 116
QY 60 PPARVTIKMECNNOYNGSNPDDCSVAKGKMWSSSDNVGNNGSYMEKH-IPPPM 118
Db 117 TNERRVIVPADPTLMTSTHROWLEMAIKEYGLMEIDTISFFQNMKGELCKNKEDFLR 176

QY 119 TNERRVIVPADPTLMTSTHROWLEMAIKEYGLPDVDILLFQNDGKELCKMTKDDFOR 178
Db 177 STSIYTEVLISHLYLRDSSSSLYG-NTQAHDTQSSRLTANEDPSYEAVRSGMGNMS 235
QY 179 LTPSYNADILLSHLYLRGATFIEPNTSVPEATQRTTRDPLPEQARRSAM-TSSH 237
Db 236 SPYTKSPMGQONVKKSDQOSQOPDPYQILGPTSSRLANPSSGQIOLMQLLELSDS 295
QY 238 HPQSKATQSSSTVPEKT-EDQRPOLDPYQILGPTSSRLANPSSGQIOLMQLLELSDS 296
Db 296 SNASCITWEGTNGEFKMTDPDEVARRMGERKSKPMNNDKLSRLRYXXDNIMTKVHGR 355
QY 297 SNASCITWEGTNGEFKMTDPDEVARRMGERKSKPMNNDKLSRLRYXXDNIMTKVHGR 356
Db 356 RYAKKDFPHGIAQALOPHPTISMRYKYPSEFSYMPYSHQOKVNPSPHSSMPTSSG 415
QY 357 RYAKKDFPHGIAQALOPHPTISMRYKYPSEFSYMPYSHQOKVNPSPHSSMPTSSG 416
Db 416 FFGATSPYMNSPANITYPNPVRHPTVHSHLGCFY 453
QY 417 FFAAPNPYMNSTPGIYPTN-T-RLPAHMPSHLGTY 451
RESULT 6
ENTRY 529844 #type complete
TITLE transforming protein fl1, short splice form - human
ALTERNATE_NAMES ERGB Friend leukemia integration protein 1; transcription factor
ORGANISM #formal_name Homo sapiens #common_name man
DATE 02-Dec-1993 #sequence_revision 23-May-1997 #text_change 02-Sep-1997
ACCESSIONS S29844
REFERENCE S29844
#authors Hromas, R.; May, W.; Denny, C.; Raskind, W.; Moore, J.; Maki,
#journal R.A.; Beck, E.; Klemasz, M.J.
#title Blochim. Biophys. Acta (1993) 1172:155-158
#cross_references Human FLI-1 localizes to chromosome 11Q24 and has an aberrant transcript in neuroepithelioma.
#accession S29844
#status preliminary; nucleic acid sequence not shown;
translation not shown
#molecule_type mRNA
#residues 1-385 #label HRO
#cross_references EMBL:M93255; NID:g18265; PID:g182661
#note the nucleotide sequence was submitted to the EMBL Data Library, May 1992
GENETICS
#gene GDB:FLI1
#cross_references GDB:127565; OMIM:193067
#map_position 11q24.1-11q24.3
CLASSIFICATION #superfamily transcription factor erg; ets DNA-binding domain
KEYWORDS homology; ets RII regulatory region homology
alternative splicing; DNA binding; nucleus; proto-oncogene;
transcription factor; transforming protein
FEATURE
52-126 #domain ets RII regulatory region homology #label ETS2\N
217-295 #domain ets DNA-binding domain homology #label ETS
SUMMARY #length 385 #molecular_weight 43437 #checksum 4889
Query Match 60.0% Score 1973; DB 1; Length 385;
Best Local Similarity 67.8%; Pred. No. 0.00e+00;
Matches 257; Conservative 68; Mismatches 46; Indels 8; Gaps 6;
Db 10 ARSPDSCSVSKSKLVGGESNPMYNSYMDKKNPPPPNMTNRRVIVPADATLVNQ 69
QY 78 SRNSPDDCSVAKGKMWSSSDNVGNNGSYMEKH-IPPPM TNERRVIVPADPTLMTST 136
Db 70 EHVROWLEMAIKEYGLMEIDTISFFQNMKGELCKNKEDFLRATTLTYNTEVLISHLSYLR 129
QY 137 DHVROWLEMAIKEYGLPDVDILLFQNDGKELCKMTKDDFORLTPSYNADILLSHLYLR 196
Db 130 E-SSLLAY-NTSHTDQSSRLVAKEDPSYDVRGAMNNMNGLKSPPLGGAQTISKN 187

Oy	197	ENGATFFIFPNTSVPEAOTRITTRPLPLEQARNSAMTSHSPTOSKATPPSSS-TYPKT	255
Db	188	TEORPORDEPYILGPTSSRLANPGSGOIQLOMFLLELLSDSANASCITWEGNGEFKMTD	247
Oy	256	EDQROLDPYIQLGPTSSRLANPGSGOIQLOMFLLELLSDSSNSNCITWEGNGEFKMTD	315
Db	248	PDEVARRMGQRKSRRPMNMDKLSRALRYYYDKIMTKVHKGRAYAKFDFIGIAOLAPHP	307
Oy	316	PDEVARRMGERSKRPMMNDKLSRALRYYYDKIMTKVHKGRAYAKFDFIGIAOLAPHP	375
Db	308	TSSSYKKYSDDSIYSMSYSAHOOKVNFVPHPSPSSMPVTSSFFGAASQXYTS-TGGITYPN	366
Oy	376	PSSSMTKYPSDLPIYMSSTYAHPOKNMFVAPHPALPVSSSFPAAPNPYNSTPGTIYPN	435
Db	367	PNVPRPHNTVPSHLGSY 385	
Oy	436	T--RLPAAHMPSHLGTY 451	
SOUT	7	A54617 #type fragment	
ENTRY		transcription factor erg - mouse (fragment)	
ALTERNATE_NAMES		Ig heavy chain enhancer-binding protein	
ORGANISM		#normal_name Mus musculus #common_name house mouse	
DATE		25-Oct-1994 #sequence_revision 18-Nov-1994 #text_change 30-May-1997	
REFERENCES		A54617	
AUTHORS		Rivera, R.R.; Stultey, M.H.; Steenbergen, R.; Murre, C.	
JOURNAL		Mol. Cell. Biol. (1993) 13:7163-7169	
CITE		Ets proteins: new factors that regulate immunoglobulin heavy-chain gene expression.	
#cross-references		MUID:94019387	
#accession		A54617	
#status		preliminary	
#molecule_type		mRNA	
#residues		1-272 #label RIV	
#experimental_source		pre-B-cell line 22D6	
#note		sequence extracted from NCBI backbone (NCBIN:138524), NCBIP:138524)	
CLASSIFICATION		#superfamily transcription factor erg; ets DNA-binding domain homology; ets RII regulatory region homology DNA binding; phosphoprotein; proto-oncogene; transcription factor; transforming protein	
KEYWORDS			
FEATURE			
140-218		#domain ets DNA-binding domain homology #label ETS	
SUMMARY		#length 272 #checksum 3754	
Query Match		46.7%; Score 1535; DB 2; length 272;	
Best Local Similarity		94.5%; Pred. No. 8,79e-296;	
Matches		208; Conservative 9; Mismatches 2; Indels 1; Gaps 1;	
Ddb	53	GAAFFPNTSVPEAOTRITTRPDLYEPFRRSAMTGHSILTROSKRAQSPSPAVPTED	112
Oy	199	GAFFFPNTSVPEAOTRITTRPDLYEQARRSAMTSHSPT-QSKATQSSSSIVPTKD	257
Ddb	113	QRPOLDPYIILGPTSSRLANPGSGOIQLOMFLLELLSDSSNSNCITWEGNGEFKMTDP	172
Oy	258	QRPOLDPYIILGPTSSRLANPGSGOIQLOMFLLELLSDSSNSNCITWEGNGEFKMTDP	317
Ddb	173	EVARRMGERKSRRPMNYDKLSRALRYYYDKIMTKVHKGRAYAKFDFIGIAOLAPHP	232
Oy	318	EVARRMGERKSRRPMNYDKLSRALRYYYDKIMTKVHKGRAYAKFDFIGIAOLAPHP	377
Ddb	233	SSLKYPSDLPTWGSTHAHPOKMNFVSPHPALPVSSSF 272	
Oy	378	SMYKYPDDLPIYMSYSAHPOKNMFVAPHPALPVSSSF 417	
RESULT	8	A56646 #type fragment	
ENTRY		transcription factor erg/fli-1 homolog - sea urchin	
TITLE			

ORGANISM		(Lytechinus variegatus) (fragment)
DATE	03-Oct-1995	#organism_name Lytechinus variegatus #common_name variegated urchin
ACCESSIONS	08-Sep-1997	#sequence_revision 03-Oct-1995 #text_change
REFERENCE	A56546	
#authors	QI, S.; Chen, Z.Q.; Pappas, T.S.; Lautenberger, J.A.	
#journal	DNA Seq. (1992) 3:127-130	
#title	The sea urchin erg homolog defines a highly conserved erg-specific domain.	
#accession	A56546	
##status	preliminary	
##molecule_type	DNA	
##residues	1-173	##label QIA
##cross-references	GB:M81067; NID:g161310; PID:g161311	
CLASSIFICATION	#superfamily transcription factor erg; ets DNA-binding domain homology; ets RII regulatory region homology	
KEYWORDS	DNA binding; nucleus; transcription factor	
FEATURE		
6-84	#domain ets DNA-binding domain homology #label ETS	
SUMMARY	#length 173 #checksum 5090	
Query Match	29.2%	Score 959; DB 2; Length 173;
Best Local Similarity	73.3%	Pred. No. 6.95e-171;
Matches	129; Conservative	27; Mismatches 12; Indels 8; Gaps 6;
Db	1	GSGIOLOMFLLELSDSNASHITTEGNGEKLVDPEETARRMEKRSKPNNYDKLSR 60
Qy	280	SGOIQLOMOFLLELSDSSNSNCITWGCTGGEFFMTPTDDEYARRMGERRSKPPNNYDKLSR 339
Db	61	ALRRYYDNKNIMTKVHGKRYAYKFDFAGLAQAMOPVQADPSPMYRXODLYLPGRH--PTR 118
Qy	340	ALRRYYDKNIMTKVHGKRYAYKFDFAGILAQADPHPPSESMYKXPSPDLPMSSVHAHPQK 399
Db	119	LNFVG-I-PIN-EPTNASLFSSHSTYSSTPGTGANITYSGRHTHHASHMSHIGTY 172
Qy	400	MNFVAHPHALPVLTSSSFPAAPNPYNPSPTGG-IYPNTRL--P-AAHMPSHLGTTY 451
RESULT	9	
ENTRY	S51226	#type fragment
TITLE	transcription factor erg/fil-1 homology - polychaete (Nereis diversicolor) (fragment)	
ORGANISM	#formal_name Nereis diversicolor #common_name sandworm	
DATE	15-Jul-1995 #sequence_revision 23-May-1997 #text_change	
ACCESSIONS	S51226	
REFERENCE	S51226	
#authors	Lelievre-Chotteau, A.; Laudet, V.; Flourens, A.; Begue, A.; Leprince, D.; Fontaine, F.	
#journal	FEMS Lett. (1994) 354:62-66	
#title	Identification of two ets related genes in a marine worm, the polychaete annelid Nereis diversicolor.	
#accession	S51226	
##status	preliminary	
##molecule_type	DNA	
##residues	1-179	##label LEL
CLASSIFICATION	#superfamily transcription factor erg; ets DNA-binding domain homology; ets RII regulatory region homology	
KEYWORDS	DNA binding; nucleus; transcription factor	
FEATURE		
7-85	#domain ets DNA-binding domain homology #label ETS	
SUMMARY	#length 179 #checksum 4104	
Query Match	27.2%	Score 895; DB 2; Length 179;
Best Local Similarity	67.6%	Pred. No. 3.51e-157;
Matches	121; Conservative	32; Mismatches 17; Indels 9; Gaps 8;
Db	1	GSGIOLOMFLLELSDSNASHITTEGNGEKLVDPEETARRMEKRSKPNNYDKLS 60
Qy	279	GSGIOLOMFLLELSDSSNSNCITWEGNGEKMDPDVEARWRGERSKPPNNYDKLS 338

Db	61	RALRYYYDKNIMTKVHGKRYKAYKFFDPAGLAQAOAPSTDPAAKYQODL	LM5GX	HTS	118
Oy	339	RALRYYYDKNIMTKVHGKRYKAYKFFDPAGLAQAOAPSTDPAAKYQODL	LM5GX	HTS	118
Db	119	KNTLMAAHP-MASSASGCFPPAPYMSLVGSNTLXPINSHHMSHHPGHSSHTGSGY	176		
Oy	399	KNTLMAAHP-MASSASGCFPPAPYMSLVGSNTLXPINSHHMSHHPGHSSHTGSGY	176		
RESULT	10				
ENTRY		s37616	#type complete		
TITLE		transcription factor cly - fruit fly (Drosophila melanogaster)			
ALTERNATE_NAMES		transcription factor ets-97D			
ORGANISM		formal_name Drosophila melanogaster			
DATE		13-Jan-1995 #sequence_revision 23-May-1997 #text_change 18-Sep-1998			
ACCESSIONS		s37616, s28822			
REFERENCE		s37616			
#authors		The, S.M.; Xie, X.; Smyth, F.; Papas, T.S.; Watson, D.K.; Schulz, R.A.			
#journal		Oncogene (1992) 7:2471-2478			
#title		Molecular characterization and structural organization of D-cly, an ets proto-oncogene-related gene of Drosophila.			
#accession		s37616			
#status		preliminary			
#molecule_type		DNA			
#residues		1-464 #label THE			
#cross-references		EMBL:X68259; NID:g7942; PID:g7943			
REFERENCE		s28819			
#authors		Chen, T.; Bunting, M.; Karim, F.D.; Thummel, C.S.			
#journal		Dev. Biol. (1992) 151:176-191			
#title		Isolation and characterization of five Drosophila genes that encode an ets-related DNA binding domain.			
#accession		s28822			
#molecule_type		RNA			
#residues		298-449 #label CHE			
#cross-references		EMBL:M88477; NID:9157189; PID:9552088			
#note		the authors translated the codon AGC for residue 302 as Thr			
GENETICS					
#gene		FlyBase:ET597D			
#cross-references		FlyBase:F8gn0004510			
#introns		22/1; 60/2; 83/3; 322/1			
CLASSIFICATION		superfamily transcription factor elg; ets DNA-binding domain			
KEYWORDS		homology; ets RII regulatory region homology			
FEATURE		DNA binding; nucleus; transcription factor			
190-263		#domain ets RII regulatory region homology #label ERR\			
348-426		#domain ets DNA-binding domain homology #label ETS			
SUMMARY		#length 464 #molecular-weight 52658 #checksum 2714			
Query Match		20.18; Score 660; DB 1; Length 464;			
Best Local Similarity		39.68; Pred. NO.2.17e-107;			
Matches		97; Conservative 62; Mismatches 74; Indels 12; Gaps 11;			
Db	192	EOIRLKIPAAANSMTHAHYVYLEMAVKOFELGINSNDWQ-NGGOELCATHEEFNOKL	250		
Oy	121	NERRAVVPADPLTMSIDVHQMELMAVKEGLPVDILLQONIDKRELCKMTKDF-QRL	179		
Db	251	-PDDGNIFMTHLQJLKE--CNFV-SVYHRAED-QRKPQPRIMASNSISTN-SGGSLS	304		
Oy	180	TESYNADILLSHHYRERGATFIFPNTSYVPEATORITTRDLPYEQARSAMTSHSP	239		
Db	305	LEGRIRKRSQSV-KSSD-SVE-STISSMPSNVTITGSGNNGVOLMOPILLELTPOEH	361		
Oy	240	TOSKATQPSSTYPTKEDQRPQDLPYQIOLPDS-SRLANGSGQIQLOFLELLSDSN	298		
Db	362	TDVIEVWGTGEFELNDPDPVAVRAMEGKKRKPANVYKLSRALYXYDGMISVSGKRF	421		
Oy	299	SNCITTEGNGEERKMDPDEVARAMEGKKRKPANVYKLSRALYXYDGMISVSGKRY	358		
Db	422	AYKFD 426			

[illegible]

```

ALTERNATE_NAMES      GABPA: nuclear respiratory factor-2 alpha chain;
                      transcription factor E4TF1 60K chain
ORGANISM              #formal_name Homo sapiens #common_name man
DATE                  21-Jan-1994 #sequence_revision 23-May-1997 #text_change
                      05-Jun-1998
ACCESSIONS            A48146; A46303; A55903
REFERENCE              A48146
#authors              Watanabe, H.; Sawade, J.; Yano, K.; Yamaguchi, K.; Goto, M.;
#journal              Mol. Cell. Biol. (1993) 13:1385-1391
#title                CDNA Cloning of transcription factor E4TF1 subunits with Ets
                      and notch motifs.
#cross-references      MIMD:93180783
#accession            A48146
#status               not compared with conceptual translation
#molecule_type       mRNA; protein
#residues             1-454 #label WAT
#note                 sequence extracted from NCBI backbone (NCBIP:125762)
                      parts of this sequence were determined by protein
                      sequencing
REFERENCE              A46303
#authors              Vithasius, J.V.; Vithasius, C.A.; Scarpulla, R.C.
#journal              Genes Dev. (1993) 7:380-392
#title                Identity of GABP with NRF-2, a multisubunit activator of
                      cytochrome oxidase expression, reveals a cellular role for
                      an ETS domain activator of viral promoters.
#accession            A46303
#molecule_type       protein
#residues             237-248;429-448 #label VIR
REFERENCE              A55903
#authors              Cardot, P.; Pastier, D.; Lacorte, J.M.; Mangeney, M.; Zannis
                      V.I.; Chambaz, J.
#journal              Biochemistry (1994) 33:12139-12148
#title                Purification and characterization of nuclear factors binding
                      to the negative regulatory element D of human
                      apolipoprotein A-II promoter: a negative regulatory effect
                      is reversed by GABP, an ets-related protein.
#accession            A55903
#molecule_type       protein
#residues             26-34,'A',36,'X',38,'X',40-46;151-165;167-174;350-359
                      #label CAR
#label CAR
GENETICS              GDB:GABPA; E4TF1-60; E4TF1A
#gene                 #cross-references GDB:138476; OMIM:600609
#map_position         21q21-21q22.1
COMPLEX               GA-binding protein is a heterotetramer of two alpha and two
                      beta-type chains.
FUNCTION              #description a transcription factor that binds (via the alpha chain) to
                      GA-rich promoters
                      #pathway known to promote transcription of apolipoprotein A-II,
                      cytochrom c oxidase chain IV, ATPase beta chain, and
                      adenovirus E4 genes
CLASSIFICATION        #superfamily transcription factor elg; ets DNA-binding domain
                      homology; ets RII regulatory region homology
                      DNA binding; nucleus; transcription factor
KEYWORDS              #domain ets RII regulatory region homology #label ERR\
                      #domain ets DNA-binding domain homology #label ETS
FEATURE               174-245
                      322-400
SUMMARY               #length 454 #molecular_weight 51311 #checksum 3079
                      19.8%; Score 649; DB 1; Length 454;
                      Best Local Similarity 41.6%; Pred. No. 4,38e-105;
                      Matches 101; Conservative 60; Mismatches 62; Indels 20; Gaps 15;
Db 176 EORGLGIPYPIOWSTQVLTLMVYVWVKFSSMTDIDITLT-NISGRELCSLNQEDFFQV 234
OY 121 NERRIVPADETLWSTSHVROMEMAYKEXGLDPVDITLTFQNDIGKELCKMTKDDF-QRL 179
Db 235 -PR--GEILMHELRL-K--YYLAAQ--QQMNE-IVTI-QQP-VQIIPAS-VQSARP 281
OY 180 TPSTNADLLSHLYLKERGATFFFPNTSVPEATQATITTRPDLPEQARRSAMTSHSP 239

```

Db	282	TTIKVIN--SSAAKAAVVO--RAPRISGEDRSSP--GNFRGN--NGOIQLOMFLELLTPTDKAR	336
Oy	240	TOSKATQPSSTSVPTEDORPOLDPYQILGPTSSRLANGSGOIQLOMFLELLSSSNS	299
Db	337	DCISAVGDEGEFKLNOPELVAQKWGORKKPNMNTSEKLSRALRYYYDGMICKVOGKRFV	396
Oy	300	NCITMEGTNGEKKMTPDEVARRMGERSKKPMNNTDKLSRALRYYYDKNIMTVHGKRVA	359
Db	397	YKF 399	
Oy	360	YKF 362	
RESULT	13		
ENTRY		138739	#type complete
TITLE			nuclear respiratory factor-2 subunit alpha - human
ORGANISM			#formal_name Homo sapiens #common_name man
DATE			29-May-1998 #sequence_revision 29-May-1998 #text_change 10-Jul-1998
ACCESSIONS		138739	
REFERENCE		138739	
#authors			Gugneja, S.; Virbasius, J.V.; Scarpulla, R.C.
#journal			Mol. Cell. Biol. (1995) 15:102-111
#title			Four structurally distinct, non-DNA-binding subunits of human nuclear respiratory factor 2 share a conserved transcriptional activation domain.
#cross-references			MUID:95097980
#accession		138739	
#status			preliminary; translated from GB/EMBL/DBJ
#molecule_type			mRNA
#residues			1-454 #label RES
#cross-references			EMBL:U03044; NID:9531892; PID:9531893
CLASSIFICATION			#superfamily transcription factor elg; ets DNA-binding domain homology; ets RII regulatory region homology
FEATURE			
174-245			#domain ets RII regulatory region homology #label ERR
SUMMARY			#length 454 #molecular_weight 51408 #checksum 3243
Query Match			19.5%; Score 640; DB 2; Length 454;
Best Local Similarity			40.7%; Pred. No. 3.35e-103;
Matches			99; Conservative 61; Mismatches 63; Indels 20; Gaps 15;
Db	176	EOERGIPIPDLPQMSITDOYLHVVVWVKFESMTDIDTLT--NISGEELCSLNODEFFQRY	234
Oy	121	NERRYIVPADPLMTSTDHRCWLEMAVKEXGLPDVDILFONIDGKELCKMTKDDF--QRL	179
Db	235	-PR--GEILMSHLELLR-K--YVLASQE--QQMNQ--IYTI--DQ--VQIIPAS--VQSATP	281
Oy	180	TPSYADLILSHLYLRREGATFEPNPSVYEDATQRIITRPDLPEQARRSMTWSHP	239
Db	282	TTIKVINCAKA-AVVO--RAPRISGEDRSSP--GNFRGN--NGOIQLOMFLELLTPTDKAR	336
Oy	240	TOSKATQPSSTSVPTEDORPOLDPYQILGPTSSRLANGSGOIQLOMFLELLSSSNS	299
Db	337	DCISAVGDEGEFKLNOPELVAQKWGORKKPNMNTSEKLSRALRYYYDGMICKVOGKRFV	396
Oy	300	NCITMEGTNGEKKMTPDEVARRMGERSKKPMNNTDKLSRALRYYYDKNIMTVHGKRVA	359
Db	397	YKF 399	
Oy	360	YKF 362	
RESULT	14		
ENTRY		S28823	#type fragment
TITLE			transcription factor ets-21c - fruit fly (Drosophila melanogaster) (fragment)
ALTERNATE_NAMES			transforming protein ets-6
ORGANISM			#formal_name Drosophila melanogaster
DATE			31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 23-May-1997
ACCESSIONS		S28823	
REFERENCE		S28819	

Job time : 68 secs.

```

#authors      Chen, T.; Bunting, M.; Karim, F. D.; Thummel, C.S.
#journal      Dev. Biol. (1992) 151:176-191
#title        Isolation and characterization of five Drosophila genes that
               encode an ets-related DNA binding domain.
#accession    S28823
#molecule_type DNA
#residues     1-103 #label CHE
#cross-references EMBL:M88475

GENETICS
#gene         Ets21C
#cross-references FlyBase:FBgn0005660
#introns      64/2
CLASSIFICATION #superfamily transcription factor erg; ets DNA-binding domain
                homology; ets RII regulatory region homology
KEYWORDS       DNA binding; nucleus; transcription factor
FEATURE        8-86
SUMMARY        #domain ets DNA-binding domain homology #label ETS
                #length 103 #checksum 5825

Query Match      19.0%; Score 623; DB 2; Length 103;
Best Local Similarity 80.4%; Pred. No. 1.19e-99;
Matches 78; Conservative 13; Mismatches 6; Indels 0; Gaps 0;

Db      2  SGGQIQLMQFLLELADSSNANAIWEGSGEPRLDPDEVARRWGERKAKPMNMYDKLS 61
QY      279 GSGQIQLMQFLLELSDSSNSNCITWGTNGEFTKMTDPDEVARRWGERKSKPMNMYDKLS 338
Db      62  RALRYYYDKNIMTKVHGKRYAYKDFPHGLAACQACA 98
QY      339 RALRYYYDKNIMTKVHGKRYAYKDFPHGLAQLQPHR 375

RESULT 15
ENTRY   S24300 #type complete
TITLE   transforming protein (elg) - fruit fly (Drosophila
         melanogaster)
ORGANISM #formal_name Drosophila melanogaster
DATE      13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change
         16-Feb-1997
ACCESSIONS
REFERENCE S24300
          S24300
          Pribyl, L.J.; Watson, D.K.; Schulz, R.A.; Papas, T.S.
          Oncogene (1991) 6:1175-1183
          #authors
          #journal
          #title      D-elg, a member of the Drosophila ets gene family: sequence,
                     expression and evolutionary comparison.
                     S24300
#accession S24300
#status    preliminary
#molecule_type mRNA
#residues  1-211 #label PRI
#cross-references EMBL:X58481

GENETICS
#gene      FlyBase:Ets97D
#cross-references FlyBase:FBgn0004510
CLASSIFICATION #superfamily ets DNA-binding domain homology
KEYWORDS       DNA binding
FEATURE        95-173
SUMMARY        #domain ets DNA-binding domain homology #label ETS
                #length 211 #molecular_weight 23603 #checksum 6044

Query Match      15.0%; Score 493; DB 2; Length 211;
Best Local Similarity 59.8%; Pred. No. 8.80e-73;
Matches 58; Conservative 24; Mismatches 14; Indels 1; Gaps 1;

Db      77  MNPSYTTIGSGNNGOVQIQFLLELTDCEHTDVLEWGTGEFEKLTDPDEVARRWGERK 136
QY      268 LGPTS-SRLANPGSQIQLMQFLLELSDSSNSNCITWGTNGEFTKMTDPDEVARRWGER 326
Db      137  KKKPMANTEKSLRALRYYYDGMISKVSGKRPAKKFD 173
QY      327  KSKPMNMYDKLSRALRYYYDKNIMTKVHGKRYAYKFD 363

```


ID	ERG_MOUSE	STANDARD	PRT	272 AA.
Db	416	FEGATSPYNNPSANIVPNPVPHPTTHVSHLGGFV	453	: :
Qy	417	FFAANPYWNSPTGCIYPTNT---KLPAAMHSHLGTYY	451	: :
RESULT	5	STANDARD	PRT	272 AA.
ID	ERG_MOUSE	STANDARD	PRT	272 AA.
AC	P81270			
DT	15-JUL-1998 (REL. 36, CREATED)			
DT	15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)			
DT	15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)			
DE	TRANSFORMING PROTEIN ERG (FRAGMENT).			
GN	ERG OR ERG-3.			
OS	MUS MUSCULUS (MOUSE).			
OC	EUAROTIA; METAFOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;			
CC	EUTHERIA; RODENTIA.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE: 94019387.			
CC	RIVERA R. R., STUIVER M. H., STEENBERGEN R., MURRE C.;			
CC	MOL. CELL. BIOL. 13:7163-7169(1993).			
CC	-1- SUBCELLULAR LOCATION: NUCLEAR.			
CC	-1- SIMILARITY: BELONGS TO THE ETS FAMILY.			
CC	EMBL: S66169; G436277; -			
DR	MED: MGI:95415; ERG.			
DR	PROSITE: PS00345; ETS_DOMAIN.1; 1.			
DR	PROSITE: PS00346; ETS_DOMAIN.2; 1.			
DR	PROSITE: PS50061; ETS_DOMAIN.3; 1.			
KW	NUCLEAR PROTEIN; TRANSFORMING PROTEIN; DNA-BINDING; PROTO-ONCOGENE.			
FT	NON_TER 1 1			
FT	DNA_BIND 138 218			
FT	NON_TER 272 272			
SO	SEQUENCE 272 AA; 30907 MW; D2C258D9 CRC32;			
	Query Match	46.7%;	Score 1535;	DB 1; Length 272;
	Best Local Similarity	94.5%;	Pred. No. 0.00e+00;	
	Matches 208; Conservative 9; Mismatches 2; Indels 1; Gaps 1;			
Db	53	GAATFPTSVYPEATORTITRPLDLYEPYPRRSAMTGHSHLTPSKAAPSAPVPTED	112	: :
Qy	199	GATFPTSVYPEATORTITRPLDLYEQARSAWTSHTSPT-QSKATQSSSVPTED	257	: :
Db	113	QRPQLDPQIIGIPSSRLANPGSGQIQLMQLELLSDSSNSCITWGTNGEFTKMTDPD	172	: :
Qy	258	QRPQLDPQIIGIPSSRLANPGSGQIQLMQLELLSDSSNSCITWGTNGEFTKMTDPD	317	: :
Db	173	EVARRMGRKSKKPNMNVYKLSRALRYDYDKIMKIVGKRAYKEDPHGIAQLAOLPHPE	232	: :
Db	318	EVARRMGRKSKKPNMNVYKLSRALRYDYDKIMKIVGKRAYKEDPHGIAQLAOLPHPE	377	: :
Db	233	SLYKYPDDLPMGSGYHAHPQKMFVSPHPALPVTSSSF	272	: :
Qy	378	SSMYKYPDDLPMGSGYHAHPQKMFVSPHPALPVTSSSF	417	: :
RESULT	6	STANDARD	PRT	173 AA.
ID	ERG_LYTVA	STANDARD	PRT	173 AA.
AC	001414;			
DT	01-FEB-1995 (REL. 31, CREATED)			
DT	01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)			
DT	01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)			
DE	ERG PROTEIN HOMOLOG (FRAGMENT).			
GN	ERG.			
OS	LYTECHINUS VARIEGATUS (SEA URCHIN).			
OC	EUAROTIA; METAFOA; ECHINODERMATA; ECHINOZOA; ECHINOIDEA;			
CC	EUTCHINOIDEA.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE: 93091246.			
RA	QI S., CHEN Z.O., PAPAS T.S., LAUTENBERGER J.A.;			
RL	DNA SEQ. 3:127-129(1992).			
CC	-1- SUBCELLULAR LOCATION: NUCLEAR.			
CC	-1- SIMILARITY: BELONGS TO THE ETS FAMILY.			

Query Match
Best Local Similarity 73.3%, Score 959; DB 1; Length 173;
Matches 129; Conservative 27; Mismatches 12; Indels 8; Gaps 6;

Dd 1 SGOIQIOWFLELLSDSSNNANCITWEGTNGEFKTTDDEVARRKGKKSPNMYDKLSR 60
|||
Qy 280 SGOIQIOWFLELLSDSSNSNCITWEGTNGEFTKDDEVARRGEEKSKPNMYDKLSR 339
||||

Dd 61 ALRYRYOKNMTKYHGKRVAYAKPEPAGLAOMPOVADPSMRYROSLTYLPGVH--PTK 118
|||
Qy 340 ALRYRYDDNMNTKYHGKRVAKEPFHFIAQLPHPESSMKYKPSLPSSTHAHPQK 399
||||

Dd 119 LNFPGT-PIN-PTSNASLFSSHSSSYWSSPTGANIPSGHYTHPHASHMSHIGYY 172
::|:
Qy 400 MNFAHPHPALPYISSSFEAPARNYNMSPISGG-IYPNTRL-P-AAMNPGLGIYY 451
||||

RESULT 7
ID ELG_DROME STANDARD; PRT; 464 AA.
AC 004688;
DT 01-JUN-1994 (REL. 29, CREATED)
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
De DNA-BINDING PROTEIN D-ELG.
GN ETS97D OR ELG.
OS DIROSOPHILA MELANOGASTER (FRUIT FLY).
OC EUKARYOTA; METAZOA; ARTHROPODA; INSECTA; DIPTERA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 93096481.
RA THE S.M., XIE X., SMYTH F., PAPAS T.S., WATSON D.K., SCHULTZ R.A.;
RL ONOGENE 7:2471-2478(1992).
RN [2]
RP SEQUENCE OF 323-463 FROM N.A.
RX MEDLINE; 91319397.
RA PRIBYL L.J., WATSON D.K., SCHULZ R.A., PAPAS T.S.;
RL ONOGENE 6:1175-1183(1991).
RN [3]
RP SEQUENCE OF 298-449 FROM N.A.
RX MEDLINE; 92249640.
RA CHEN T., BUNTING M., KARIM F.D., THUMMEL C.S.;
RL DEV. BIOL. 151:176-191(1992).
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- SIMILARITY: BELONGS TO THE ETS FAMILY.
DR EMBL; X68259; G1943; -.
DR EMBL; X58481; G1941; ALT_INIT.
DR EMBL; M88471; G552088; -.
DR PIR; S24300; S24300.
DR PIR; S28822; S28822.
DR PIR; S37616; S37616.
DR TRANSFAC; T02085; -.
DR FLTBASE; FBgn0004510; Ets97D.
DR PROSITE; PS00345; ETS_DOMAIN_1; 1.
DR PROSITE; PS00346; ETS_DOMAIN_2; 1.
DR PROSITE; PS50061; ETS_DOMAIN_3; 1.
DNA-BINDING; NUCLEAR PROTEIN.
FT DNA_BIND 346 426 ETS-DOMAIN.
SQ SEQUENCE 464 AA; 52658 MW; A759B7A CRC32;

Query March 20.1%; Score 660; DB 1; Length 464;
Best Local Similarity 39.6%; Pred. No. 5.29e-124;
Matches 97; Conservative 62; Mismatches 74; Indels 12; Gaps 11;

Dd 192 EOIRLKIPLEANENTHAVYTWEMAVAKOEIVGINNSDMQG-MNGDELCAHTHEEFNOKL 250

QY 121 NERVIAPADPTLWSTHVRQWLEMAVKEYGLPVDILLFONIDGKELCKTKDQDF-QRL 179
Db 251 -PRRPGNIFWTHQLOLKE--CNFV-SYVHKRAEE-QRKPRKPRMSANSISTN-SGGSLS 304
QY 180 TPSYNADILLSHLYLBERGATFIFPNTSVYPEATQRTTPDLPYEQARSAMTSHSH 239
Db 305 LEQRIMRKSYSOV-KSSD-SVE-STTSSMNPSTTIGSGNNGOVOLQWOFLELLTDCHE 361
QY 240 TQSRATQPSSTVTKEDQRPQDLPYQILGPTS-SRLANPSSGOIQLOWOFLELLSDSN 298
Db 362 TDVIEWGTGEGFKLTPDDRYARLWGEKKKPPANWEXLSALRYDDGDMISVSGRRF 421
QY 299 SNTCTWEGTNGEFKMTDPEVARRMGERKSKPMNMYDLSPALRYDDKINTKVGKRY 358
Db 422 AYKFD 426
QY 359 AYKFD 363

ULT 8 GABA_MOUSE STANDARD: PRT: 454 AA.

000422;
01-DEC-1992 (REL. 24, CREATED)
01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)
15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE GA BINDING PROTEIN ALPHA CHAIN (GABP-ALPHA SUBUNIT).
GABPA.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 91343912.
RA LAMARCO K., THOMPSON C.C., BYERS B.P., WALTON E.M., MCKNIGHT S.L.;
SCIENCE 253:789-792(1991).
RN [2]
RX X-RAY CRYSTALLOGRAPHY (2.15 ANGSTROMS) OF 320-320.
RP MEDLINE: 98128030.
RA BATCHELOR A.H., PIPER D.E., DE LA BROUSSE F.C., MCKNIGHT S.L.,
WOLBERGER C.;
SCIENCE 279:1037-1041(1998).
CC -1- FUNCTION: TRANSCRIPTION FACTOR CAPABLE OF INTERACTING WITH PURINE
RICH REPEATS (GA REPEATS).
CC -1- SUBUNIT: HETEROETRAMER OF TWO ALPHA AND TWO BETA SUBUNITS.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- TISSUE SPECIFICITY: OBITUOUS.
CC -1- SIMILARITY: BELONGS TO THE ETS FAMILY.
EMBL: M74515; G193383; -.
PIR: A40858; A40858.
PDB: 1AWC; 18-MAR-98.
DR TRANSFAC: T00298; -.
DR TRANSFAC: T01402; -.
DR MGD; MGI:95610; GABPA.
DR PROSITE: PS00345; ETS_DOMAIN_1; 1.
DR PROSITE: PS00346; ETS_DOMAIN_2; 1.
DR PROSITE: PS50061; ETS_DOMAIN_3; 1.
KM TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN; 3D-STRUCTURE.
FT DNA_BIND 320 400 ETS-DOMAIN.
SQ SEQUENCE 454 AA; 51363 MW; C3BA5F76 CRC32;

Query Match 19.9% Score 653; DB 1; Length 454;
Best Local Similarity 41.6% Pred. No.2,61e-122;
Matches 101; Conservative 61; Mismatches 61; Indels 20; Gaps 15;

Db 176 EOERLGIPYDPFIRNSTDQVHLVWVWVWKEFEMTDIDLTLT-NISGRELSLNQEDFFQRY 234
QY 121 NERVIAPADPTLWSTHVRQWLEMAVKEYGLPVDILLFONIDGKELCKTKDQDF-QRL 179
Db 235 -PR--GELIWSHLELR-K--YVLASOE--QQONE-IVTI-DQP-VQIIPAS-VPPATP 281
QY 180 TPSYNADILLSHLYLBERGATFIFPNTSVYPEATQRTTPDLPYEQARSAMTSHSH 239

Db 282 TTIKVIN-SSAKAKVO-RSPRISGEDRSSP-GNRTGN--NGOIQLOWOFLELLTDRDAR 336
QY 240 TQSRATQPSSTVTKEDQRPQDLPYQILGPTSRLANPSSGOIQLOWOFLELLSDSN 299
Db 337 DCISWGDGEGFKLNOBELVAQKQKRNKPTMYEKLRLRYYYGDMICKYQGRFV 396
QY 300 NCITWEGTNGEFKMTDPEVARRMGERKSKPMNMYDLSPALRYDDKINTKVGKRYA 359
Db 397 YKF 399
QY 360 YKF 362

RESULT 9 GABA_HUMAN STANDARD: PRT: 454 AA.

AC 006546; 012939;
01-FEB-1995 (REL. 31, CREATED)
01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE GA BINDING PROTEIN ALPHA CHAIN (GABP-ALPHA SUBUNIT) (TRANSCRIPTION
FACTOR EATF1-60) (NUCLEAR RESPIRATORY FACTOR-2 SUBUNIT ALPHA).
GN EATF1 OR GABPA.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 93180783.
RA WATANABE H., SAWADA J.-I., YANO K.-I., YAMAGUCHI K., GOTO M.,
HANDA H.;
MOL. CELL. BIOL. 13:1385-1391(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE: 95097980.
RA GUGNEVA S., VIRBASJUS J.V., SCARPUOLA R.C.;
MOL. CELL. BIOL. 15:102-111(1995).
CC -1- FUNCTION: TRANSCRIPTION FACTOR CAPABLE OF INTERACTING WITH PURINE
RICH REPEATS (GA REPEATS). NECESSARY FOR THE EXPRESSION OF THE
ADENOVIRUS EA GENE.
CC -1- SUBUNIT: HETEROETRAMER OF TWO ALPHA AND TWO BETA SUBUNITS.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- SIMILARITY: BELONGS TO THE ETS FAMILY.
DR EMBL: D13318; G286027; -.
DR EMBL: U13044; G531893; -.
DR TRANSFAC: T01390; -.
DR MIM; 600609; -.
DR PROSITE: PS00345; ETS_DOMAIN_1; 1.
DR PROSITE: PS00346; ETS_DOMAIN_2; 1.
DR PROSITE: PS50061; ETS_DOMAIN_3; 1.
KM TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN.
FT DNA_BIND 320 400 ETS-DOMAIN.
FT CONFLICT 289 290 SS -> RC (IN REF. 2).
FT CONFLICT 440 440 A -> V (IN REF. 2).
SQ SEQUENCE 454 AA; 51295 MW; E0EA85D4 CRC32;

Query Match 19.8% Score 649; DB 1; Length 454;
Best Local Similarity 41.6% Pred. No.2,42e-121;
Matches 101; Conservative 60; Mismatches 62; Indels 20; Gaps 15;

Db 176 EOERLGIPYDPFIRNSTDQVHLVWVWVWKEFEMTDIDLTLT-NISGRELSLNQEDFFQRY 234
QY 121 NERVIAPADPTLWSTHVRQWLEMAVKEYGLPVDILLFONIDGKELCKTKDQDF-QRL 179
Db 235 -PR--GELIWSHLELR-K--YVLASOE--QQONE-IVTI-DQP-VQIIPAS-VQSATP 281
QY 180 TPSYNADILLSHLYLBERGATFIFPNTSVYPEATQRTTPDLPYEQARSAMTSHSH 239
Db 282 TTIKVIN-SSAKAKVO-RAPRISGEDRSSP-GNRTGN--NGOIQLOWOFLELLTDRDAR 336
QY 240 TQSRATQPSSTVTKEDQRPQDLPYQILGPTSRLANPSSGOIQLOWOFLELLSDSN 299
Db 337 DCISWGDGEGFKLNOBELVAQKQKRNKPTMYEKLRLRYYYGDMICKYQGRFV 396

[illegible]

Similarity	59.08;	Pred. No.	2.33e-83;
Conservative	20;	Mismatches	23;
		Indels	0;
		Gaps	0;

